

One-Way Analysis of Variance

Nathaniel E. Helwig

Associate Professor of Psychology and Statistics
University of Minnesota



October 17, 2020

Copyright © 2020 by Nathaniel E. Helwig

Table of Contents

1. Model Form
2. Parameter Estimation
3. Sums of Squares
4. Omnibus F Test
5. Multiple Comparisons
 - Overview
 - Familywise Error Rate
 - Methods to Control FWER
 - Confidence Intervals

Table of Contents

1. Model Form
2. Parameter Estimation
3. Sums of Squares
4. Omnibus F Test
5. Multiple Comparisons
 - Overview
 - Familywise Error Rate
 - Methods to Control FWER
 - Confidence Intervals

One-Way Analysis of Variance Model

The one-way ANalysis Of VAriance (ANOVA) model extends the independent samples t test problem to the situation with $a > 2$ groups.

Suppose that we have an independent sample of observations where $y_{ij} \stackrel{\text{iid}}{\sim} N(\mu_j, \sigma^2)$ for $i = 1, \dots, n_j$ and $j = 1, \dots, a$.

The one-way ANOVA model assumes that

$$y_{ij} = \mu_j + \epsilon_{ij}$$

where μ_j is the mean for the j -th group, and $\epsilon_{ij} \stackrel{\text{iid}}{\sim} N(0, \sigma^2)$.

Cell Means versus Treatment Effects

The previous form of the model is known as the “cell means” model form, given that it explicitly writes each group’s mean μ_j .

We could also write the model in the “treatment effects” model form, which decomposes the group’s means such as

$$y_{ij} = \mu + \tau_j + \epsilon_{ij}$$

where μ is the overall mean that is common to all observations, and τ_j is the j -th group’s treatment effect, which satisfies $\sum_{j=1}^a \tau_j = 0$.

Note that $\mu_j = \mu + \tau_j$, which reveals that $\tau_j = \mu_j - \mu$ is the j -th group’s deviation from the overall mean.

Table of Contents

1. Model Form
2. Parameter Estimation
3. Sums of Squares
4. Omnibus F Test
5. Multiple Comparisons
 - Overview
 - Familywise Error Rate
 - Methods to Control FWER
 - Confidence Intervals

Least Squares Estimation

To estimate the μ_j parameters, consider minimizing the least squares loss function

$$\sum_{j=1}^a \sum_{i=1}^{n_j} (y_{ij} - \mu_j)^2 = \sum_{j=1}^a L_j(\mu_j | \mathbf{y}_j)$$

where $L_j(\cdot)$ is the least squares loss function for the j -th group, and $\mathbf{y}_j = (y_{1j}, \dots, y_{n_jj})^\top$ is the vector of observed data for the j -th group.

The least squares loss function for each group can be written as

$$L_j(\mu_j | \mathbf{y}_j) = \sum_{i=1}^{n_j} (y_{ij} - \mu_j)^2 = \sum_{i=1}^{n_j} y_{ij}^2 - 2\mu_j n_j \bar{y}_j + n_j \mu_j^2$$

where $\bar{y}_j = \frac{1}{n_j} \sum_{i=1}^{n_j} y_{ij}$ is the j -th group's sample mean.

Least Squares Solution

Taking the derivative of $L_j(\mu_j|\mathbf{y}_j)$ with respect to μ_j , setting the derivative to zero, and solving for μ_j reveals that

$$\hat{\mu}_j = \bar{y}_j$$

is the least squares estimate of the j -th group's mean.

For the treatment effects parameterization, the overall mean estimate is

$$\hat{\mu} = \frac{1}{a} \sum_{j=1}^a \hat{\mu}_j = \frac{1}{a} \sum_{j=1}^a \bar{y}_j$$

and the treatment effects are estimated using

$$\hat{\tau}_j = \hat{\mu}_j - \hat{\mu}$$

Table of Contents

1. Model Form
2. Parameter Estimation
- 3. Sums of Squares**
4. Omnibus F Test
5. Multiple Comparisons
 - Overview
 - Familywise Error Rate
 - Methods to Control FWER
 - Confidence Intervals

Some Insight to the Model Name

Why do we call it an “analysis of *variance*” model when we are analyzing mean differences in the data?

The model is referred to as an “analysis of variance” model because the model makes it possible to analyze how the variance in the data can be decomposed into the within and between group variances.

As we shall see, it is possible to decompose...

- the sample variance into within and between group variances
- the corresponding degrees of freedom into within and between DoF

Sum of Squares Total

The *sum of squares total* is defined as

$$\text{SST} = \sum_{j=1}^a \sum_{i=1}^{n_j} (y_{ij} - \bar{y})^2$$

where $\bar{y} = \frac{1}{n} \sum_{j=1}^a \sum_{i=1}^{n_j} y_{ij}$ is the overall mean and $n = \sum_{j=1}^a n_j$ is the total sample size.

Note that if we multiplied the SST quantity by $\frac{1}{n-1}$, this would be the sample variance of the data ignoring the group membership.

The SST quantity has $n - 1$ degrees of freedom, given that there are n independent values and one constraint: $\sum_{j=1}^a \sum_{i=1}^{n_j} (y_{ij} - \bar{y}) = 0$.

Sum of Squares Between

The *sum of square between* groups is defined as

$$\text{SSB} = \sum_{j=1}^a \sum_{i=1}^{n_j} (\bar{y}_j - \bar{y})^2 = \sum_{j=1}^a n_j (\bar{y}_j - \bar{y})^2$$

which shows how the group means \bar{y}_j vary around the overall mean \bar{y} .

Note that if $n_j = n_k$ for all j, k , then the SSB quantity is proportional to the sample variance of the group means.

The SSB quantity has $a - 1$ degrees of freedom, given that there are a independent group means and one constraint: $\sum_{j=1}^a n_j (\bar{y}_j - \bar{y}) = 0$.

Sum of Squares Within

Finally, the *sum of squares within* groups is defined as

$$\text{SSW} = \sum_{j=1}^a \sum_{i=1}^{n_j} (y_{ij} - \bar{y}_j)^2$$

which shows how the data vary around the group means \bar{y}_j .

If we multiplied the SSW quantity by $\frac{1}{n-1}$ this would be the pooled estimate of the variance conditioned on the group membership.

The SSW has $n - a$ degrees of freedom because there are n independent values and a constraints: $\sum_{i=1}^{n_j} (y_{ij} - \bar{y}_j) = 0$ for $j = 1, \dots, a$.

Decomposing the SS and DoF

The degrees of freedom of the three sums of square are additive, i.e.,

$$df_T = df_B + df_W$$

where $df_T = n - 1$ is the SST degrees of freedom, $df_B = a - 1$ is the SSB degrees of freedom, and $df_W = n - a$ is the SSW degrees of freedom.

Although it may not be obvious from their definitions, the sums of squares can be decomposed such as

$$SST = SSB + SSW$$

which implies that the total variation (SST) is the summation of the between group variation (SSB) and the within group variation (SSW).

Proof of SS Decomposition

To prove the relationship, we can write

$$\begin{aligned} \text{SST} &= \sum_{j=1}^a \sum_{i=1}^{n_j} (y_{ij} - \bar{y}_j + \bar{y}_j - \bar{y})^2 \\ &= \sum_{j=1}^a \sum_{i=1}^{n_j} (\bar{y}_j - \bar{y})^2 + \sum_{j=1}^a \sum_{i=1}^{n_j} (y_{ij} - \bar{y}_j)^2 + 2 \sum_{j=1}^a \sum_{i=1}^{n_j} (y_{ij} - \bar{y}_j) (\bar{y}_j - \bar{y}) \end{aligned}$$

where the first term is the SSB, the second term is the SSW, and the third term is zero:

$$\begin{aligned} \sum_{j=1}^a \sum_{i=1}^{n_j} (y_{ij} - \bar{y}_j) (\bar{y}_j - \bar{y}) &= \sum_{j=1}^a \sum_{i=1}^{n_j} (y_{ij} \bar{y}_j - y_{ij} \bar{y} - \bar{y}_j^2 + \bar{y}_j \bar{y}) \\ &= \sum_{j=1}^a n_j \bar{y}_j^2 - n \bar{y}^2 - \sum_{j=1}^a n_j \bar{y}_j^2 + n \bar{y}^2 = 0 \end{aligned}$$

R-squared and Adjusted R-squared

The *coefficient of multiple determination* (aka R-squared) is defined as

$$R^2 = \frac{SSB}{SST} = 1 - \frac{SSE}{SST}$$

which is the proportion of the total variation in the sample of data that can be accounted for by the group membership.

The adjusted coefficient of multiple determination (or adjusted R-squared) value is defined as

$$R_{adj}^2 = 1 - \frac{SSW/df_W}{SST/df_T} = 1 - \frac{\hat{\sigma}^2}{s_y^2}$$

where $\hat{\sigma}^2$ is the mean squared error (defined in the next section) and s_y^2 is the sample variance of Y ignoring the group membership.

Table of Contents

1. Model Form
2. Parameter Estimation
3. Sums of Squares
4. Omnibus F Test
5. Multiple Comparisons
 - Overview
 - Familywise Error Rate
 - Methods to Control FWER
 - Confidence Intervals

Are Any Means Different?

Suppose that we want to test whether there are any meant differences.

This corresponds to testing the null hypothesis $H_0 : \mu_j = \mu \ \forall j$ versus the alternative hypothesis $H_1 : (\exists j)(\mu_j \neq \mu)$.

- Same as testing $H_0 : \tau_j = 0 \ \forall j$ versus $H_1 : (\exists j)(\tau_j \neq 0)$

To test the null hypothesis, we can use the F test statistic

$$F = \frac{SSB/df_B}{SSW/df_W}$$

where the numerator is the *mean square between* (or treatment) and the denominator is the *mean square within* (or error).

ANOVA Table

We typically organize the SS information into an ANOVA table:

Source	SS	df	MS	F	p-value
SSB	$\sum_{j=1}^a \sum_{i=1}^{n_j} (\bar{y}_j - \bar{y})^2$	$a - 1$	MSB	F	p
SSW	$\sum_{j=1}^a \sum_{i=1}^{n_j} (y_{ij} - \bar{y}_j)^2$	$n - a$	MSW		
SST	$\sum_{j=1}^a \sum_{i=1}^{n_j} (y_{ij} - \bar{y})^2$	$n - 1$			
$MSB = \frac{SSB}{a-1}$, $MSW = \frac{SSW}{n-a}$, and $F = \frac{MSB}{MSW}$					

Assuming that H_0 is true, $F \sim F(df_B, df_W) = F(a - 1, n - a)$

The p-value for testing H_0 is given by $p = P(F_{a-1, n-a} > F)$

Estimating the Error Variance

The mean square within is an unbiased estimate of the error variance, i.e., the quantity

$$\hat{\sigma}^2 = \text{SSW}/df_W$$

has expectation $E(\hat{\sigma}^2) = \sigma^2$, where σ^2 is the variance of the ϵ terms in the ANOVA model.

The unbiasedness of $\hat{\sigma}^2$ can be proved using a similar approach that was used in the Parameter Estimation notes to show that the sample variance (dividing by $n - 1$) is unbiased.

The Limitation of the Omnibus Test

This is referred to as an *omnibus* test because this approach simultaneously tests the significance of all of the treatment effects.

Note that if we reject the null hypothesis, this does not inform us which of the group means significantly differ from one another.

As a result, it is necessary to use a different approach if your goal is to determine which means are significantly different from one another.

Memory Example: Data Description

Visual and auditory cues example from Hays (1994) *Statistics*.

- Does lack of visual/auditory synchrony affect memory?

Total of $n = 30$ college students participate in memory experiment.

- Watch video of person reciting 50 words
- Try to remember the 50 words (record number correct)

Randomly assign $n_j = 10$ subjects to one of $g = 3$ video conditions:

- **fast**: sound precedes lip movements in video
- **normal**: sound synced with lip movements in video
- **slow**: lip movements in video precede sound

Memory Example: Descriptive Statistics

Number of correctly remembered words (y_{ij}):

Subject (i)	Fast ($j = 1$)	Normal ($j = 2$)	Slow ($j = 3$)
1	23	27	23
2	22	28	24
3	18	33	21
4	15	19	25
5	29	25	19
6	30	29	24
7	23	36	22
8	16	30	17
9	19	26	20
10	17	21	23
$\sum_{i=1}^{10} y_{ij}$	212	274	218
$\sum_{i=1}^{10} y_{ij}^2$	4738	7742	4810

Memory Example: OLS Estimation (by hand)

The least-squares estimates of μ_j are the sample means:

$$\hat{\mu}_1 = \bar{y}_1 = \frac{1}{10} \sum_{i=1}^{10} y_{i1} = 212/10 = 21.2$$

$$\hat{\mu}_2 = \bar{y}_2 = \frac{1}{10} \sum_{i=1}^{10} y_{i2} = 274/10 = 27.4$$

$$\hat{\mu}_3 = \bar{y}_3 = \frac{1}{10} \sum_{i=1}^{10} y_{i3} = 218/10 = 21.8$$

Memory Example: Input Data (in R)

```
> words <- c(23, 27, 23, 22, 28, 24, 18, 33, 21, 15,
+           19, 25, 29, 25, 19, 30, 29, 24, 23, 36,
+           22, 16, 30, 17, 19, 26, 20, 17, 21, 23)
> cond <- factor(rep(c("fast", "normal", "slow"), 10))
> words
[1] 23 27 23 22 28 24 18 33 21 15 19 25 29 25 19 30 29 24 23 36 22 16
[23] 30 17 19 26 20 17 21 23
> cond
[1] fast   normal slow   fast   normal slow   fast   normal slow
[10] fast   normal slow   fast   normal slow   fast   normal slow
[19] fast   normal slow   fast   normal slow   fast   normal slow
[28] fast   normal slow
Levels: fast normal slow
```

Memory Example: Fit Model (in R)

```
> # fit anova model
> amod <- aov(words ~ cond)
> amod
Call:
  aov(formula = words ~ cond)
```

Terms:

	cond	Residuals
Sum of Squares	233.8667	535.6000
Deg. of Freedom	2	27

Residual standard error: 4.453879

Estimated effects may be unbalanced

```
> amod$coefficients
(Intercept)  condnormal  condslow
          21.2           6.2           0.6
```

Memory Example: Omnibus Test (in R)

```
> # omnibus F test
> anova(amod)
Analysis of Variance Table

Response: words
          Df Sum Sq Mean Sq F value    Pr(>F)
cond         2  233.87  116.933   5.8947 0.007513 **
Residuals  27  535.60   19.837
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Note: we reject H_0 but we don't know which means are different!

Table of Contents

1. Model Form
2. Parameter Estimation
3. Sums of Squares
4. Omnibus F Test

5. Multiple Comparisons

- Overview

- Familywise Error Rate

- Methods to Control FWER

- Confidence Intervals

Forming Linear Combinations

To understand the nature of the mean differences, we need to perform followup tests (multiple comparisons) to determine where the mean differences are occurring in the data.

In most cases, we are interested in follow-up tests about linear combinations of the factor level means.

A linear combination of the factor level means has the form

$$L = \sum_{j=1}^a c_j \mu_j$$

where c_1, \dots, c_a are the coefficients that define the linear combination. A contrast is a linear combination where $\sum_{j=1}^a c_j = 0$.

Estimating Linear Combinations

Of course, in practice, we don't know the true population means, so the linear combination can be estimated such as

$$\hat{L} = \sum_{j=1}^a c_j \hat{\mu}_j$$

where $\hat{\mu}_j = \bar{y}_j$ are the estimated factor level means.

Note that the variance of the linear combination has the form

$$\text{Var}(\hat{L}) = \sum_{j=1}^a c_j^2 \text{Var}(\hat{\mu}_j) = \sigma^2 \sum_{j=1}^a \frac{c_j^2}{n_j}$$

which is due to the fact that $\text{Var}(\hat{\mu}_j) = \sigma^2/n_j$.

Testing Hypotheses about Linear Combinations

Suppose we want to test a null hypothesis of the form $H_0 : L = L_0$ versus the alternative hypothesis $H_1 : L \neq L_0$.

We can use a t test statistic

$$T_0 = \frac{\hat{L} - L_0}{\sqrt{\widehat{\text{Var}}(\hat{L})}}$$

where $\widehat{\text{Var}}(\hat{L}) = \hat{\sigma}^2 \sum_{j=1}^a c_j^2/n_j$ uses the MSW to estimate σ^2 .

Assuming the null hypothesis $H_0 : L = L_0$ is true, the test statistic T_0 follows a Student's t distribution with $n - a$ degrees of freedom.

Overview of the Problem

If we only wanted to test a single linear combination, we could use a Student's t test. However, when there are $a > 2$ groups, we often want to test more than a single hypothesis.

For $m \geq 2$ tests, the family-wise type I error rate (FWER) is the probability of making at least one type I error among the m tests.

As a reminder, for a single significance test, the type I error rate is

$$\alpha = P(\text{Reject } H_0 \mid H_0 \text{ is true})$$

which is the probability of a false positive (i.e., rejecting a true null).

FWER for Independent and Dependent Tests

Suppose we have m significance tests H_{0k} for $k = 1, \dots, m$, and assume that we use a significance level of α for each test.

If the m tests are independent of one another, then the FWER is

$$\text{FWER} = 1 - (1 - \alpha)^m$$

where $(1 - \alpha)^m$ is the probability of not making any type I error, which is simply the product of not making a type I error for each test.

For example...

- with $m = 3$ independent tests, we have $\text{FWER} = 0.142625$
- with $m = 5$ independent tests, we have $\text{FWER} = 0.2262191$

FWER for Dependent Tests

In practice the m tests will not necessarily (or even likely) be independent of one another.

When the tests are dependent on one another, there is no simple formula for calculating the FWER.

This is because, when the tests are dependent, the FWER will depend on the dependence structure of the m tests—which is often unknown.

So what do we do to control the FWER in general cases?

Bonferroni Correction

Suppose that we have m tests, and let α_k denote the type I error rate for the k -th test (for $k = 1, \dots, m$). Typically $\alpha_k = \alpha_0 \forall k$.

According to Boole's inequality, we know that

$$\text{FWER} \leq \sum_{k=1}^m P(\text{Reject } H_{0k} \mid H_{0k} \text{ is true}) = \sum_{k=1}^m \alpha_k$$

and this can be further simplified to $\text{FWER} = m\alpha_0$ if $\alpha_k = \alpha_0 \forall k$.

Bonferroni's solution to the multiple comparison problem is to set the α_k values such that $\sum_{k=1}^m \alpha_k = \alpha$, where α is the desired FWER.

- If $\alpha_k = \alpha_0 \forall k$, Bonferroni's correction simply sets $\alpha_0 = \alpha/m$

Tukey's Method

Tukey considered the problem of testing all of the $a(a-1)/2$ possible pairwise comparisons of the form $H_0 : \mu_j - \mu_k = 0 \forall j, k$ versus the alternative $H_1 : (\exists j, k)(\mu_j \neq \mu_k)$.

The variance of any pairwise comparison has the form

$$\text{Var}(\hat{L}_{jk}) = \sigma^2 \left(\frac{1}{n_j} + \frac{1}{n_k} \right)$$

which simplifies to $\text{Var}(\hat{L}_{jk}) = \sigma^2(2/n_0)$ if $n_j = n_k = n_0$.

To test the (j, k) -th pairwise comparison, Tukey assumed that $n_j = n_k = n_0 \forall j, k$, and proposed using the test statistic

$$q = \frac{\sqrt{2}\hat{L}_{jk}}{\sqrt{\widehat{\text{Var}}(\hat{L}_{jk})}} = \frac{\bar{y}_j - \bar{y}_k}{\hat{\sigma}/\sqrt{n_0}}$$

Tukey's Method (continued)

Tukey showed that the test statistic q follows the *Studentized range distribution* with degrees of freedom $(a, n - a)$ if H_0 is true.

For balanced designs (i.e., when $n_j = n_k = n_0 \forall j, k$), using Tukey's procedure will exactly control the FWER across all $a(a - 1)/2$ possible pairwise comparisons.

When the design is unbalanced (i.e., when $n_j \neq n_k$ for some j, k), the procedure is known as the Tukey-Kramer method, and the FWER will be below (but not exactly equal to) the significance level α .

Scheffé's Method

Scheffé considered the problem of testing the significance of all possible contrasts of the factor level means.

$H_0 : \sum_{j=1}^a c_j \mu_j = 0$ for all coefficients satisfying $\sum_{j=1}^a c_j = 0$, versus
 $H_1 : \sum_{j=1}^a c_j \mu_j \neq 0$ for at least one set of coefficients with $\sum_{j=1}^a c_j = 0$.

Note that all contrasts are equal to zero if and only if $\mu_j = \mu \forall j$.

- $\mu_j = \mu \forall j$ implies that $\sum_{j=1}^a c_j \mu_j = \mu \sum_{j=1}^a c_j = 0$
- $\sum_{j=1}^a c_j \mu_j = 0$ for all contrasts means $\mu_j - \mu_k = 0 \forall j, k$

Can use the omnibus F test that we previously discussed for testing the overall significance of the ANOVA model.

Confidence Interval for a Single Linear Combination

To form a $100(1 - \alpha)\%$ confidence interval for a generic linear combination $L = \sum_{j=1}^a c_j \mu_j$, we can use the general formula

$$\hat{L} \pm C_\alpha \sqrt{\widehat{\text{Var}}(\hat{L})}$$

where

- $\hat{L} = \sum_{j=1}^a c_j \hat{\mu}_j = \sum_{j=1}^a c_j \bar{y}_j$ is the estimated linear combination
- $\widehat{\text{Var}}(\hat{L}) = \hat{\sigma}^2 \sum_{j=1}^a c_j^2 / n_j$ is the estimated variance of \hat{L}
- C_α denotes some critical value depending on the confidence level α

When forming a confidence interval for a single linear combination, the critical value is defined as $C_\alpha = t_{n-a}^{(1-\alpha/2)}$, where $t_{n-a}^{(1-\alpha/2)}$.

Confidence Intervals for Linear Combinations

For each possible method to control the FWER, the critical value would be defined as:

- None: $C_\alpha = t_{n-a}^{(1-\alpha/2)}$
- Bonferroni: $C_\alpha = t_{n-a}^{(1-\alpha/(2m))}$
- Tukey: $C_\alpha = q_{a,n-a}^{(1-\alpha)}/\sqrt{2}$
- Scheffé: $C_\alpha = \sqrt{(a-1)F_{a-1,n-a}^{(1-\alpha)}}$

When choosing a method, note that...

- Bonferroni is applicable to any m tests.
- Tukey's method is only applicable for pairwise comparisons
- Scheffé's method is only applicable to contrasts

Memory Example: Pairwise Comparisons

Revisiting the previous example, suppose that we want to form confidence intervals for each of the possible pairwise comparisons.

In this case, we could choose between all three FWER correction methods: Bonferroni, Tukey, and Scheffé.

Of course, we know that Tukey's method should be preferred for this situation, but we will still compare the three methods.

The variance of any pairwise comparison $L_{jk} = \mu_j - \mu_k$ is given by $\text{Var}(L_{jk}) = \sigma^2(2/n_0)$.

Memory Example: Critical Values

Plugging in the mean squared error for the unknown σ^2 , we have that

$$\widehat{\text{Var}}(L_{jk}) = \hat{\sigma}^2(2/n_0) = (535.6/27)(2/10) = 1071.2/270 = 3.967407$$

For a 95% confidence interval, the different critical values that we could use are given by:

- None: $C_\alpha = t_{27}^{(0.975)} = 2.051831$
- Bonferroni: $C_\alpha = t_{27}^{(1-0.05/6)} = 2.552459$
- Tukey: $C_\alpha = q_{3,27}^{(0.95)} / \sqrt{2} = 2.479418$
- Scheffé: $C_\alpha = \sqrt{2F_{2,27}^{(0.95)}} = 2.590031$

Tukey is preferred because it produces the narrowest confidence interval that will control the FWER.

Memory Example: Confidence Intervals by Hand

The three confidence intervals using Tukey's method are:

- $\hat{L}_{12} \pm (q_{3,27}^{(0.95)} / \sqrt{2}) \sqrt{\widehat{\text{Var}}(L_{jk})} =$
 $(27.4 - 21.2) \pm (2.479418) \sqrt{3.967407} = [1.2614; 11.1386]$
- $\hat{L}_{13} \pm (q_{3,27}^{(0.95)} / \sqrt{2}) \sqrt{\widehat{\text{Var}}(L_{jk})} =$
 $(21.8 - 21.2) \pm (2.479418) \sqrt{3.967407} = [-4.3386; 5.5386]$
- $\hat{L}_{23} \pm (q_{3,27}^{(0.95)} / \sqrt{2}) \sqrt{\widehat{\text{Var}}(L_{jk})} =$
 $(21.8 - 27.4) \pm (2.479418) \sqrt{3.967407} = [-10.5386; -0.6614]$

Memory Example: Confidence Intervals in R

```
> TukeyHSD(amod)
  Tukey multiple comparisons of means
    95% family-wise confidence level
```

```
Fit: aov(formula = words ~ cond)
```

```
$cond
```

	diff	lwr	upr	p adj
normal-fast	6.2	1.261409	11.1385914	0.0117228
slow-fast	0.6	-4.338591	5.5385914	0.9513012
slow-normal	-5.6	-10.538591	-0.6614086	0.0238550