

What is the complete guide for interpreting ANOVA results in R?

Authored by
stats writer

April 29, 2024

RECOMMENDED CITATION

stats writer (2024). *What is the complete guide for interpreting ANOVA results in R?*. PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=141212>

The complete guide for interpreting ANOVA results in R is a comprehensive resource that outlines the necessary steps for understanding and analyzing the outcome of an ANOVA test in the R programming language. This guide covers the basics of ANOVA, including its purpose and assumptions, as well as the necessary statistical concepts and techniques for interpreting the results. It also includes detailed instructions on how to conduct ANOVA tests in R, as well as how to interpret and communicate the results in a meaningful way. This guide is a valuable tool for anyone looking to gain a thorough understanding of ANOVA and its application in R for data analysis and interpretation.

Complete Guide: Interpret ANOVA Results in R

A one-way ANOVA is used to determine whether or not there is a statistically significant difference between the means of three or more independent groups.

This tutorial provides a complete guide on how to interpret the results of a one-way ANOVA in R.

Step 1: Create the Data

Suppose we want to determine if three different workout programs lead to different average weight loss in individuals.

To test this, we recruit 90 people to participate in an experiment in which we randomly assign 30 people to follow either program A, program B, or program C for one month.

The following code creates the data frame we'll be working with:

```
#make this example reproducible
```

```
set.seed(0)
```

```
#create data frame
```

```
data <- data.frame(program = rep(c('A', 'B', 'C'), each =  
30),
```

```
weight_loss = c(runif(30, 0, 3),
```

```
runif(30, 0, 5),
```

```
runif(30, 1, 7)))
```

```
#view first six rows of data frame
```

```
head(data)
```

```
program weight_loss
```

```
1 A 2.6900916
```

```
2 A 0.7965260
```

```
3 A 1.1163717
```

```
4 A 1.7185601
```

```
5 A 2.7246234
```

```
6 A 0.6050458
```

Step 2: Perform the ANOVA

Next, we'll use the `aov()` command to perform a one-way ANOVA:

```
#fit one-way ANOVA model
```

```
model <- aov(weight_loss ~ program, data = data)
```

Step 3: Interpret the ANOVA Results

Next, we'll use the `summary()` command to view the results of the one-way ANOVA:

```
#view summary of one-way ANOVA model
```

```
summary(model)
```

```
Df Sum Sq Mean Sq F value Pr(>F)
program 2 98.93 49.46 30.83 7.55e-11 ***
Residuals 87 139.57 1.60
```

```
---
```

```
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Here's how to interpret every value in the output:

Df program: The degrees of freedom for the variable *program*. This is calculated as `#groups - 1`. In this case, there were 3 different workout programs, so this value is: $3 - 1 = 2$.

Df Residuals: The degrees of freedom for the residuals. This is calculated as $\# \text{total observations} - \# \text{ groups}$. In this case, there were 90 observations and 3 groups, so this value is: $90 - 3 = 87$.

Sum Sq program: The sum of squares associated with the variable *program*. This value is 98.93.

Mean Sq. Program: The mean sum of squares associated with program. This is calculated as $\text{Sum Sq. program} / \text{Df program}$. In this case, this is calculated as: $98.93 / 2 = 49.46$.

Mean Sq. Residuals: The mean sum of squares associated with the residuals. This is calculated as $\text{Sum Sq. residuals} / \text{Df residuals}$. In this case, this is calculated as: $139.57 / 87 = 1.60$.

F Value: The overall F-statistic of the ANOVA model. This is calculated as $\text{Mean Sq. program} / \text{Mean sq. Residuals}$. In this case, it is calculated as: $49.46 / 1.60 = 30.83$.

Pr(>F): The p-value associated with the F-statistic with numerator df = 2 and denominator df = 87. In this case,

the p-value is 7.552e-11, which is an extremely tiny number.

The most important value in the entire output is the p-value because this tells us whether there is a significant difference in the mean values between the three groups.

Recall that a one-way ANOVA uses the following null and alternative hypotheses:

H0 (null hypothesis): All group means are equal.

HA (alternative hypothesis): At least one group mean is different from the rest.

Since the p-value in our ANOVA table (7.552e-11) is less than .05, we have sufficient evidence to reject the null hypothesis.

This means we have sufficient evidence to say that the mean weight loss experienced by the individuals is not equal between the three workout programs.

Step 4: Perform Post-Hoc Tests (If Necessary)

If the p-value in the ANOVA output is less than .05, we reject the null hypothesis. This tells us that the mean

value between each group is not equal. However, it doesn't tell us *which* groups differ from each other.

In order to find this out, we must perform a . In R, we can use the TukeyHSD() function to do so:

```
#perform Tukey post-hoc test
TukeyHSD(model)

$program
diff lwr upr p adj
B-A 0.9777414 0.1979466 1.757536 0.0100545
C-A 2.5454024 1.7656076 3.325197 0.0000000
C-B 1.5676610 0.7878662 2.347456 0.0000199
```

Here's how to interpret the results:

The adjusted p-value for the mean difference between group A and B is .0100545. The adjusted p-value for the mean difference between group A and C is .0000000. The adjusted p-value for the mean difference between group B and C is .0000199.

Since each of the adjusted p-values is less than .05, we can conclude that there is a significant difference in

mean weight loss between *each* group.

ARABPSYCHOLOGY.COM