



A menu icon consisting of three horizontal bars of increasing length from left to right.

🕒 DECEMBER 3, 2021 🚩 BY ZACH

## How to Perform Post-Hoc Pairwise Comparisons 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.

A one-way ANOVA uses the following null and alternative hypotheses:

- $H_0$ : All group means are equal.
- $H_A$ : Not all group means are equal.

If the overall **p-value** of the ANOVA is less than a certain significance level (e.g.  $\alpha = .05$ ) then we reject the null hypothesis and conclude that not all of the group means are equal.

In order to find out which group means are different, we can then perform **post-hoc pairwise comparisons**.

The following example shows how to perform the following post-hoc pairwise comparisons in R:

- The Tukey Method
- The Scheffe Method

- The Bonferroni Method
- The Holm Method

## Example: One-Way ANOVA in R

Suppose a teacher wants to know whether or not three different studying techniques lead to different exam scores among students. To test this, she **randomly assigns** 10 students to use each studying technique and records their exam scores.

We can use the following code in R to perform a one-way ANOVA to test for differences in mean exam scores between the three groups:

```
#create data frame
df <- data.frame(technique = rep(c("tech1", "tech2", "tech3"), each=10),
                 score = c(76, 77, 77, 81, 82, 82, 83, 84, 85, 89,
                           81, 82, 83, 83, 83, 84, 87, 90, 92, 93,
                           77, 78, 79, 88, 89, 90, 91, 95, 95, 98))

#perform one-way ANOVA
model <- aov(score ~ technique, data = df)

#view output of ANOVA
summary(model)

Df Sum Sq Mean Sq F value Pr(>F)
technique     2   211.5   105.73   3.415 0.0476 *
Residuals    27   836.0    30.96
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

The overall p-value of the ANOVA (.0476) is less than  $\alpha = .05$  so we'll reject the null hypothesis that the mean exam score is the same for each studying technique.

We can proceed to perform post-hoc pairwise comparisons to determine which groups have different means.

## The Tukey Method

The Tukey post-hoc method is best to use when the sample size of each group is equal.

We can use the built-in **TukeyHSD()** function to perform the Tukey post-hoc method in R:

```
#perform the Tukey post-hoc method
TukeyHSD(model, conf.level=.95)

Tukey multiple comparisons of means
95% family-wise confidence level

Fit: aov(formula = score ~ technique, data = df)

$technique
      diff      lwr      upr      p adj
tech2-techn1 4.2 -1.9700112 10.370011 0.2281369
tech3-techn1 6.4  0.2299888 12.570011 0.0409017
tech3-techn2 2.2 -3.9700112  8.370011 0.6547756
```

From the output we can see that the only p-value (“**p adj**”) less than .05 is for the difference between technique and technique 3.

Thus, we would conclude that there is only a statistically significant difference in mean exam scores between students who used technique 1 and technique 3.

## The Scheffe Method

The Scheffe method is the most conservative post-hoc pairwise comparison method and produces the widest confidence intervals when comparing group means.

We can use the **ScheffeTest()** function from the **DescTools** package to perform the Scheffe post-hoc method in R:

```
library(DescTools)

#perform the Scheffe post-hoc method
ScheffeTest(model)

Posthoc multiple comparisons of means: Scheffe Test
95% family-wise confidence level

$technique
      diff      lwr.ci      upr.ci     pval
tech2-tech1 4.2 -2.24527202 10.645272 0.2582
tech3-tech1 6.4 -0.04527202 12.845272 0.0519 .
tech3-tech2 2.2 -4.24527202  8.645272 0.6803

---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 156
```

From the output we can see that there are no p-values less than .05, so we would conclude that there is no statistically significant difference in mean exam scores among any groups.

## The Bonferroni Method

The Bonferroni method is best to use when you have a set of planned pairwise comparisons you'd like to make.

We can use the following syntax in R to perform the Bonferroni post-hoc method:

```
#perform the Bonferroni post-hoc method
pairwise.t.test(df$score, df$technique, p.adj='bonferroni')
```

Pairwise comparisons using t tests with pooled SD

data: df\$score and df\$technique

tech1 tech2

tech2 0.309 -

tech3 0.048 1.000

P value adjustment method: bonferroni

From the output we can see that the only p-value less than .05 is for the difference between technique and technique 3.

Thus, we would conclude that there is only a statistically significant difference in mean exam scores between students who used technique 1 and technique 3.

## The Holm Method

The Holm method is also used when you have a set of planned pairwise comparisons you'd like to make beforehand and it tends to have even higher power than the Bonferroni method, so it's often preferred.

We can use the following syntax in R to perform the Holm post-hoc method:

```
#perform the Holm post-hoc method
pairwise.t.test(df$score, df$technique, p.adj='holm')
```

Pairwise comparisons using t tests with pooled SD

```
data: df$score and df$technique
```

```
tech1 tech2  
tech2 0.206 -  
tech3 0.048 0.384
```

```
P value adjustment method: holm
```

From the output we can see that the only p-value less than .05 is for the difference between technique and technique 3.

Thus, again we would conclude that there is only a statistically significant difference in mean exam scores between students who used technique 1 and technique 3.

## Additional Resources

The following tutorials provide additional information about ANOVA's and post-hoc tests:

[How to Interpret the F-Value and P-Value in ANOVA](#)

[The Complete Guide: How to Report ANOVA Results](#)

[Tukey vs. Bonferroni vs. Scheffe: Which Test Should You Use?](#)



Published by Zach

[View all posts by Zach](#)

## PREV

[How to Concatenate Strings in R \(With Examples\)](#)