How can a three-way ANOVA be performed in R?

Authored by

stats writer

July 1, 2024

A three-way ANOVA (analysis of variance) is a statistical test used to determine the impact of three independent variables on a dependent variable. In order to perform a three-way ANOVA in R, the following steps can be followed:

1. Load the necessary packages: The first step is to load the necessary packages in R, such as "stats" and "car".

2. Import the data: The next step is to import the data that will be used for the analysis into R. This can be done using the "read.csv()" function.

3. Create the ANOVA model: Using the "aov()" function, the ANOVA model can be created by specifying the dependent variable and the three independent variables.

4. Check for assumptions: Before performing the ANOVA, it is important to check for the assumptions of normality, homogeneity of variances, and independence. This can be done using diagnostic plots and statistical tests.

5. Perform the ANOVA: Once the assumptions have been met, the ANOVA can be performed using the "summary()" function. This will provide the results of the ANOVA, including the F-statistic, p-value, and effect sizes.

6. Interpret the results: The results of the ANOVA can be interpreted by looking at the significance of the main effects and interactions between the independent variables.

In conclusion, a three-way ANOVA can be performed in R by loading the necessary packages, importing the data, creating the ANOVA model, checking for assumptions, performing the ANOVA, and interpreting the results.

Perform a Three-Way ANOVA in R

A three-way ANOVA is used to determine whether or not there is a statistically significant difference between the means of three or more independent groups that have been split on three factors.

The following example shows how to perform a three-

way ANOVA in R.

Example: Three-Way ANOVA in R

Suppose a researcher wants to determine if two training programs lead to different mean improvements in jumping height among college basketball players.

The researcher suspects that gender and division (Division I or II) may also affect jumping height so he collects data for these factors as well.

His goal is to perform a three-way ANOVA to determine how training program, gender, and division affect jumping height.

Use the following steps to perform this three-way ANOVA in R:

Step 1: Create the Data

First, let's create a data frame to hold the data:

```
#create dataset
df <- data.frame(program=rep(c(1, 2), each=20),
gender=rep(c('M', 'F'), each=10, times=2),
division=rep(c(1, 2), each=5, times=4),
```

```
height=c(7, 7, 8, 8, 7, 6, 6, 5, 6, 5,
5, 5, 4, 5, 4, 3, 3, 4, 3, 3,
6, 6, 5, 4, 5, 4, 5, 4, 4, 3,
2, 2, 1, 4, 4, 2, 1, 1, 2, 1))

#view first six rows of dataset
head(df)

program gender division height
1 1 M 1 7
2 1 M 1 7
3 1 M 1 8
4 1 M 1 8
5 1 M 1 7
6 1 M 2 6
```

Step 2: View Descriptive Statistics

Before performing the three-way ANOVA, we can use to quickly summarize the mean jumping height increase grouped by training program, gender, and divison:

```
library(dplyr)

#calculate mean jumping height increase grouped by
```

```
program, gender, and division
df %>%
group_by(program, gender, division) %>%
summarize(mean_height = mean(height))

# A tibble: 8 x 4
# Groups: program, gender
program gender division mean_height

1 1 F 1 4.6
2 1 F 2 3.2
3 1 M 1 7.4
4 1 M 2 5.6
5 2 F 1 2.6
6 2 F 2 1.4
7 2 M 1 5.2
8 2 M 2 4
```

Here's how to interpret the output:

The mean jumping height increase among division I females who used training program 1 was 4.6 inches.The mean jumping height increase among division II females who used training program 1 was 3.2 inches.The mean jumping height increase among

division I males who used training program 1 was 7.4 inches.

And so on.

Step 3: Perform the Three-Way ANOVA

```
#perform three-way ANOVA
model <- aov(height ~ program * gender * division, data=df)

#view summary of three-way ANOVA
summary(model)
```

```
Df Sum Sq Mean Sq F value Pr(>F)
program 1 36.1 36.10 65.636 2.98e-09 ***
gender 1 67.6 67.60 122.909 1.71e-12 ***
division 1 19.6 19.60 35.636 1.19e-06 ***
program:gender 1 0.0 0.00 0.000 1.000
program:division 1 0.4 0.40 0.727 0.400
gender:division 1 0.1 0.10 0.182 0.673
program:gender:division 1 0.1 0.10 0.182 0.673
Residuals 32 17.6 0.55
---
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

The Pr(>F) column shows the p-value for each individual factor and the interactions between the factors.

From the output we can see that none of the interactions between the three factors were statistically significant.

We can also see that each of the three factors - program, gender, and division - were statistically significant.

We can now use dplyr once again to find the mean jumping height increase for program, gender, and division separately:

```
library(dplyr)

#find mean jumping increase by program
df %>%
group_by(program) %>%
summarize(mean_height = mean(height))

# A tibble: 2 x 2
program mean_height
```

```
1 1 5.2
2 2 3.3

#find mean jumping increase by gender
df %>%
group_by(gender) %>%
summarize(mean_height = mean(height))

# A tibble: 2 x 2
gender mean_height

1 F 2.95
2 M 5.55

#find mean jumping increase by division
df %>%
group_by(division) %>%
summarize(mean_height = mean(height))

# A tibble: 2 x 2
division mean_height

1 1 4.95
2 2 3.55
```

From the output we can observe the following:

The mean jumping height increase among individuals who used training program 1 (5.2 inches) was higher than the mean increase among individuals who used training program 2 (3.3 inches).The mean jumping height increase among males (5.55 inches) was higher than the mean increase among females (2.95 inches).The mean jumping height increase among division 1 players (4.95 inches) was higher than the mean increase among division 2 players (3.55 inches).

In conclusion, we would state that training program, gender, and division are all significant predictors of the jumping height increase among players.

We would also state that there are no significant interaction effects between these three factors.

Additional Resources

The following tutorials explain how to fit other ANOVA models in R: