

How do you perform a nested ANOVA in R step-by-step?

Authored by
stats writer

April 27, 2024

RECOMMENDED CITATION

stats writer (2024). *How do you perform a nested ANOVA in R step-by-step?*. PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=140252>

Nested ANOVA (Analysis of Variance) is a statistical method used to analyze the effects of multiple categorical independent variables on a continuous dependent variable. It is commonly used in research studies to determine whether there are significant differences between groups or conditions.

To perform a nested ANOVA in R, follow these steps:

Step 1: Organize your data

- Ensure that your data is in a spreadsheet format, with the dependent variable in one column and the independent variables in separate columns.
- If your data is not already in this format, use the "reshape" function in R to restructure it.

Step 2: Load the necessary packages

- In order to perform nested ANOVA in R, you will need to load the "car" and "lme4" packages.

Step 3: Fit the nested model

- Use the "aov" function to fit the nested model, specifying the dependent variable and the nested independent variables.
- For example, if you have two independent variables (A and B) nested within a third variable (C), the syntax would be: `model <- aov(dependent variable ~ A/B + Error(C/A), data = dataset)`

Step 4: Check the model assumptions

- Use the "plot" function to check for normality, homogeneity of variances, and independence of residuals.
- If these assumptions are not met, consider transforming the data or using a different statistical test.

Step 5: Conduct the ANOVA test

- Use the "anova" function to conduct the nested ANOVA test, which will provide information about the significance of each independent variable and the interaction between them.

Step 6: Interpret the results

- Look at the p-values from the ANOVA table to determine which variables have a significant effect on the dependent variable.
- If the overall model is significant, you can use post-hoc tests (such as Tukey's HSD) to determine which groups or conditions are significantly different from each other.

In conclusion, nested ANOVA in R is a useful tool for analyzing the effects of multiple independent variables on a continuous dependent variable. By following these steps, you can perform this statistical analysis and interpret the results accurately.

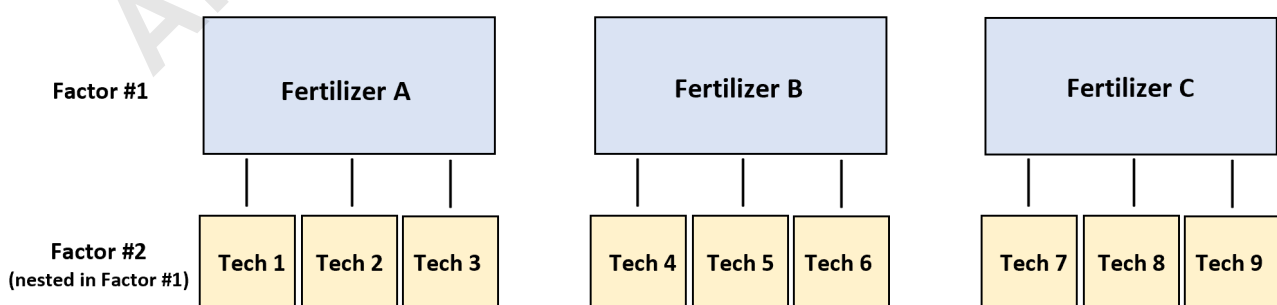
Perform a Nested ANOVA in R (Step-by-Step)

A is a type of ANOVA ("analysis of variance") in which at least one factor is *nested* inside another factor.

For example, suppose a researcher wants to know if three different fertilizers produce different levels of plant growth.

To test this, he has three different technicians sprinkle fertilizer A on four plants each, another three technicians sprinkle fertilizer B on four plants each, and another three technicians sprinkle fertilizer C on four plants each.

In this scenario, the is plant growth and the two factors are technician and fertilizer. It turns out that technician is *nested* within fertilizer:



The following step-by-step example shows how to

perform this nested ANOVA in R.

Step 1: Create the Data

First, let's create a data frame to hold our data in R:

```
#create data
```

```
df <- data.frame(growth=c(13, 16, 16, 12, 15, 16, 19, 16,  
15, 15, 12, 15,  
19, 19, 20, 22, 23, 18, 16, 18, 19, 20, 21, 21,  
21, 23, 24, 22, 25, 20, 20, 22, 24, 22, 25, 26),  
fertilizer=c(rep(c('A', 'B', 'C'), each=12)),  
tech=c(rep(1:9, each=4)))
```

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

```
head(df)
```

```
growth fertilizer tech
```

```
1 13 A 1
```

```
2 16 A 1
```

```
3 16 A 1
```

```
4 12 A 1
```

```
5 15 A 2
```

```
6 16 A 2
```

Step 2: Fit the Nested ANOVA

We can use the following syntax to fit a nested ANOVA in R:

```
aov(response ~ factor A / factor B)
```

where:

response: The response variable
factor A: The first factor
factor B: The second factor nested within the first factor

The following code shows how to fit the nested ANOVA for our dataset:

```
#fit nested ANOVA  
nest <- aov(df$growth ~ df$fertilizer / factor(df$tech))  
  
#view summary of nested ANOVA  
summary(nest)
```

```
Df Sum Sq Mean Sq F value Pr(>F)  
df$fertilizer 2 372.7 186.33 53.238 4.27e-10 ***  
df$fertilizer:factor(df$tech) 6 31.8 5.31 1.516 0.211  
Residuals 27 94.5 3.50
```

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

Step 3: Interpret the Output

From the table above, we can see that fertilizer has a statistically significant effect on plant growth (p-value < .05) but technician does not (p-value = 0.211).

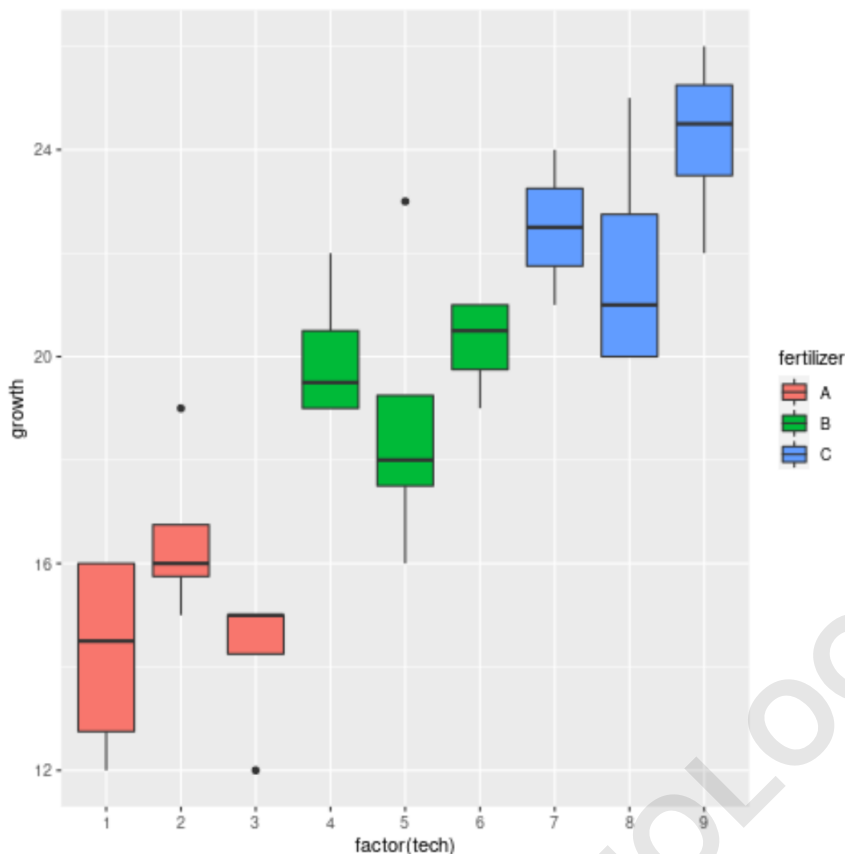
This tells us that if we'd like to increase plant growth, we should focus on the fertilizer being used rather than the individual technician who is sprinkling the fertilizer.

Step 4: Visualize the Results

Lastly, we can use boxplots to visualize the distribution of plant growth by fertilizer and by technician:

```
#load ggplot2 data visualization package  
library(ggplot2)
```

```
#create boxplots to visualize plant growth  
ggplot(df, aes(x=factor(tech), y=growth, fill=fertilizer)) +  
geom_boxplot()
```



From the chart we can see that there is significant variation in growth between the three different fertilizers, but not as much variation between the technicians within each fertilizer group.

This seems to match up with the results of the nested ANOVA and confirms that fertilizer has a significant effect on plant growth but individual technicians do not.