

How can the `linearHypothesis()` function in R be utilized?

Authored by
stats writer

June 25, 2024

RECOMMENDED CITATION

stats writer (2024). *How can the `linearHypothesis()` function in R be utilized?*.

PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=153126>

The `linearHypothesis()` function in R is used to test the significance of a linear combination of coefficients in a linear regression model. This function allows users to specify a hypothesis and perform a statistical test to determine if the linear combination of coefficients is significantly different from zero. This can be useful in determining the overall significance of a model or in comparing the effects of different variables on the outcome. The `linearHypothesis()` function can also be used to generate confidence intervals for the estimated coefficients. It is a powerful tool for analyzing and interpreting linear regression models in R.

Use the linearHypothesis() Function in R

You can use the `linearHypothesis()` function from the `car` package in R to test linear hypotheses in a specific regression model.

This function uses the following basic syntax:

```
linearHypothesis(fit, c("var1=0", "var2=0"))
```

This particular example tests if the regression coefficients `var1` and `var2` in the model called `fit` are jointly equal to zero.

The following example shows how to use this function in practice.

Example: How to Use linearHypothesis() Function in R

Suppose we have the following data frame in R that shows the number of hours spent studying, number of

practice exams taken, and final exam score for 10 students in some class:

#create data frame

```
df <- data.frame(score=c(77, 79, 84, 85, 88, 99, 95, 90, 92, 94),  
hours=c(1, 1, 2, 3, 2, 4, 4, 2, 3, 3),  
prac_exams=c(2, 4, 4, 2, 4, 5, 4, 3, 2, 1))
```

#view data frame

df

score hours prac_exams

1 77 1 2

2 79 1 4

3 84 2 4

4 85 3 2

5 88 2 4

6 99 4 5

7 95 4 4

8 90 2 3

9 92 3 2

10 94 3 1

Now suppose we would like to fit the following multiple

linear regression model in R:

Exam score = $\beta_0 + \beta_1(\text{hours}) + \beta_2(\text{practice exams})$

We can use the function to fit this model:

#fit multiple linear regression model

fit <- lm(score ~ hours + prac_exams, data=df)

#view summary of model

summary(fit)

Call:

lm(formula = score ~ hours + prac_exams, data = df)

Residuals:

Min 1Q Median 3Q Max

-5.8366 -2.0875 0.1381 2.0652 4.6381

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 72.7393 3.9455 18.436 3.42e-07 ***

hours 5.8093 1.1161 5.205 0.00125 **

prac_exams 0.3346 0.9369 0.357 0.73150

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

Residual standard error: 3.59 on 7 degrees of freedom
Multiple R-squared: 0.8004, Adjusted R-squared: 0.7434
F-statistic: 14.03 on 2 and 7 DF, p-value: 0.003553

Now suppose we would like to test if the coefficient for hours and prac_exams are both equal to zero.

We can use the linearHypothesis() function to do so:

```
library(car)
```

```
#perform hypothesis test for hours=0 and  
prac_exams=0  
linearHypothesis(fit, c("hours=0", "prac_exams=0"))
```

Linear hypothesis test

Hypothesis:

hours = 0

prac_exams = 0

Model 1: restricted model

Model 2: score ~ hours + prac_exams

Res.Df	RSS	Df	Sum of Sq	F	Pr(>F)
--------	-----	----	-----------	---	--------

1	9	452.10			
---	---	--------	--	--	--

```
2 7 90.24 2 361.86 14.035 0.003553 **
```

```
---
```

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

The hypothesis test returns the following values:

F test statistic: 14.035
p-value: .003553
H0: Both regression coefficients are equal to zero.
HA: At least one regression coefficient is not equal to zero.

Since the p-value of the test (.003553) is less than .05, we reject the null hypothesis.

In other words, we don't have sufficient evidence to say that the regression coefficients for hours and prac_exams are both equal to zero.

The following tutorials provide additional information about linear regression in R: