How can the coeftest() function be used in R?

Authored by

stats writer

June 25, 2024

The coeftest() function is a tool used in the R programming language to test the coefficients of a regression model. It allows for the calculation of confidence intervals and p-values for each coefficient, providing a statistical measure of the significance of each variable in the model. This function is commonly used in econometric and statistical analysis to assess the impact of independent variables on a dependent variable. By using the coeftest() function, researchers can better understand the relationships between variables in their data and make informed decisions based on the statistical significance of the coefficients.

Use the coeftest() Function in R

You can use the coeftest() function from the lmtest package in R to perform a t-test for each estimated coefficient in a regression model.

This function uses the following basic syntax:

coeftest(x)

where:

x: Name of the fitted regression model

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

Example: How to Use coeftest() 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, 3, 3, 2, 4, 5, 4, 3, 5, 4))

#view data frame
df
```

	score	hours	prac_exams
1	77	1	2
2	79	1	3
3	84	2	3
4	85	3	2
5	88	2	4
6	99	4	5
7	95	4	4
8	90	2	3
9	92	3	5
10	94	3	4

Now suppose we would like to fit the following multiple

linear regression model in R:

Exam score = β0 + β1(hours) + β2(practice exams)

We can use the function to fit this model:

```
#fit multiple linear regression model
fit <- lm(score ~ hours + prac_exams, data=df)
```

We can then use the coeftest() function to perform a t-test for each fitted regression coefficient in the model:

```
library(lmtest)

#perform t-test for each coefficient in model
coeftest(fit)
```

```
t test of coefficients:

             Estimate Std. Error t value  Pr(>|t|)
(Intercept) 68.40294    2.87227 23.8150 5.851e-08 ***
hours        4.19118    0.99612  4.2075  0.003998 **
prac_exams   2.69118    0.99612  2.7017  0.030566 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

The t test statistic and corresponding p-value is shown for each t-test:

Intercept: t = 23.8150, p = <0.000hours: t = 4.2075, p = .003998prac_exams: t = 2.7017, p = .030566H0: βi = 0 (the slope is equal to zero)HA: βi ? 0 (the slope is not equal to zero)

If the p-value of the t-test is less than some threshold (e.g. α = .05) then we reject the null hypothesis and conclude that there is a statistically significant relationship between the predictor variable and the response variable.

Since the p-value for each t-test is less than .05, we would conclude that each predictor variable in the model has a statistically significant relationship with the response variable.

In the context of this example, we would say that hours spent studying and number of practice exams taken are both statistically significant predictors of final exam score for students.

The following tutorials provide additional information

about linear regression in R: