

How can I calculate the confidence interval for a regression coefficient in R?

Authored by
stats writer

June 26, 2024

RECOMMENDED CITATION

stats writer (2024). *How can I calculate the confidence interval for a regression coefficient in R?*. PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=154339>

The confidence interval for a regression coefficient in R is a range of values that is likely to contain the true population coefficient with a certain level of confidence. To calculate this interval, one must first fit a linear regression model in R using the "lm" function. Then, the "confint" function can be used to obtain the lower and upper bounds of the confidence interval for the desired coefficient. The level of confidence can be specified using the "level" argument in the "confint" function. This process allows for a more accurate estimation of the regression coefficient and provides valuable information for making statistical inferences about the relationship between variables in a linear regression model.

Calculate Confidence Interval for Regression Coefficient in R

In a linear regression model, a regression coefficient tells us the average change in the associated with a one unit increase in the predictor variable.

We can use the following formula to calculate a confidence interval for a regression coefficient:

Confidence Interval for β_1 : $b_1 \pm t_{1-\alpha/2, n-2} * se(b_1)$

where:

b_1 = Regression coefficient shown in the regression table
 $t_{1-\alpha/2, n-2}$ = The t critical value for confidence level $1-\alpha$ with $n-2$ degrees of freedom where n is the total number of observations in our dataset
 $se(b_1)$ = The standard error of b_1 shown in the regression table

The following example shows how to calculate a confidence interval for a regression slope in practice.

Example: Confidence Interval for Regression Coefficient in R

Suppose we'd like to fit a simple linear regression model using hours studied as a predictor variable and exam score as a response variable for 15 students in a particular class:

Hours Studied	Exam Score
1	64
2	66
4	76
5	73
5	74
6	81
6	83
7	82
8	80
10	88
11	84
11	82
12	91
12	93
14	89

We can use the function to fit this simple linear regression model in R:

#create data frame

```
df <- data.frame(hours=c(1, 2, 4, 5, 5, 6, 6, 7, 8, 10, 11,
11, 12, 12, 14),
score=c(64, 66, 76, 73, 74, 81, 83, 82, 80, 88, 84, 82, 91,
93, 89))
```

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

```
#view model summary
summary(fit)
```

Call:

```
lm(formula = score ~ hours, data = df)
```

Residuals:

```
Min 1Q Median 3Q Max
```

```
-5.140 -3.219 -1.193 2.816 5.772
```

Coefficients:

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

```
(Intercept) 65.334 2.106 31.023 1.41e-13 ***
```

```
hours 1.982 0.248 7.995 2.25e-06 ***
```

```
---
```

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

```
Residual standard error: 3.641 on 13 degrees of
```

freedom

Multiple R-squared: 0.831, Adjusted R-squared: 0.818

F-statistic: 63.91 on 1 and 13 DF, p-value: 2.253e-06

Using the coefficient estimates in the output, we can write the fitted simple linear regression model as:

Score = 65.334 + 1.982*(Hours Studied)

Notice that the regression coefficient for hours is 1.982.

This tells us that each additional one hour increase in studying is associated with an average increase of 1.982 in exam score.

We can use the `confint()` function to calculate a 95% confidence interval for the regression coefficient:

#calculate confidence interval for regression coefficient for 'hours'

`confint(fit, 'hours', level=0.95)`

2.5 % 97.5 %

hours 1.446682 2.518068

Since this confidence interval doesn't contain the value 0, we can conclude that there is a statistically significant association between hours studied and exam score.

We can also confirm this is correct by calculating the 95% confidence interval for the regression coefficient by hand:

95% C.I. for β_1 : $b_1 \pm t_{1-\alpha/2, n-2} * se(b_1)$
95% C.I. for β_1 : $1.982 \pm t_{.975, 15-2} * .248$
95% C.I. for β_1 : $1.982 \pm 2.1604 * .248$
95% C.I. for β_1 :

The 95% confidence interval for the regression coefficient is .

Note #1: We used the to find the t critical value that corresponds to a 95% confidence level with 13 degrees of freedom.

Note #2: To calculate a confidence interval with a different confidence level, simply change the value for the level argument in the confint() function.

The following tutorials provide additional information

about linear regression in R:

ARABPSYCHOLOGY.COM