

How can I use a regression model in R to predict a single value?

Authored by
stats writer

April 28, 2024

RECOMMENDED CITATION

stats writer (2024). *How can I use a regression model in R to predict a single value?*. PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=140337>

A regression model in R is a statistical tool that allows us to make predictions based on the relationship between a dependent variable and one or more independent variables. To use a regression model in R to predict a single value, we first need to collect data on our dependent and independent variables. Then, we can use R's built-in functions or packages to fit the data to a regression model. Once the model is fitted, it can be used to predict a single value by inputting the values of the independent variables. The model will then use its learned relationship between the variables to estimate the value of the dependent variable. This process can be repeated for multiple single value predictions or for a range of values, providing insights and predictions for future data points.

Predict a Single Value Using a Regression Model in R

To fit a linear regression model in R, we can use the `lm()` function, which uses the following syntax:

```
model <- lm(y ~ x1 + x2, data=df)
```

We can then use the following syntax to use the model to predict a single value:

```
predict(model, newdata = new)
```

The following examples show how to predict a single value using fitted regression models in R.

Example 1: Predict Using a Simple Linear Regression Model

The following code shows how to fit a simple linear regression model in R:

```
#create data
```

```
df <- data.frame(x=c(3, 4, 4, 5, 5, 6, 7, 8, 11, 12),  
y=c(22, 24, 24, 25, 25, 27, 29, 31, 32, 36))
```

```
#fit simple linear regression model
```

```
model <- lm(y ~ x, data=df)
```

And we can use the following code to predict the response value for a new observation:

```
#define new observation
```

```
new <- data.frame(x=c(5))
```

```
#use the fitted model to predict the value for the new  
observation
```

```
predict(model, newdata = new)
```

```
1
```

```
25.36364
```

The model predicts that this new observation will have a response value of 25.36364.

Example 2: Predict Using a Multiple Linear Regression Model

The following code shows how to fit a multiple linear

regression model in R:

#create data

```
df <- data.frame(x1=c(3, 4, 4, 5, 5, 6, 7, 8, 11, 12),  
x2=c(6, 6, 7, 7, 8, 9, 11, 13, 14, 14),  
y=c(22, 24, 24, 25, 25, 27, 29, 31, 32, 36))
```

#fit multiple linear regression model

```
model <- lm(y ~ x1 + x2, data=df)
```

And we can use the following code to predict the response value for a new observation:

#define new observation

```
new <- data.frame(x1=c(5),  
x2=c(10))
```

#use the fitted model to predict the value for the new observation

```
predict(model, newdata = new)
```

1

26.17073

The model predicts that this new observation will have

a response value of 26.17073.

Potential Errors with Predicting New Values

The most common error you may run into when attempting to predict a new value is when the dataset you used to fit the regression model does not have the same column names as the new observation you're attempting to predict.

For example, suppose we fit the following multiple linear regression model in R:

```
#create data
```

```
df <- data.frame(x1=c(3, 4, 4, 5, 5, 6, 7, 8, 11, 12),  
x2=c(6, 6, 7, 7, 8, 9, 11, 13, 14, 14),  
y=c(22, 24, 24, 25, 25, 27, 29, 31, 32, 36))
```

```
#fit multiple linear regression model
```

```
model <- lm(y ~ x1 + x2, data=df)
```

Then suppose we attempt to use the model to predict the response value for this new observation:

```
#define new observation
```

```
new <- data.frame(x_1=c(5),
```

```
x_2=c(10))
```

```
#use the fitted model to predict the value for the new  
observation
```

```
predict(model, newdata = new)
```

```
Error in eval(predvars, data, env) : object 'x1' not found
```

We received an error because the column names for the new observation (x_1, x_2) do not match the column names of the original data frame (x1, x2) we used to fit the regression model.