

How do I perform a Likelihood Ratio Test in R?

Authored by
stats writer

April 23, 2024

RECOMMENDED CITATION

stats writer (2024). *How do I perform a Likelihood Ratio Test in R?*. PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=138323>

A Likelihood Ratio Test (LRT) is a statistical method used to compare the fit of two nested models in order to determine which model is most appropriate for a given set of data. In R, the LRT can be performed by using the "lrtest" function from the "lmtest" package. This function takes in the two models to be compared and calculates the likelihood ratio statistic, along with its associated p-value, to determine if there is a significant difference in fit between the two models. By performing the LRT in R, users can easily determine which model provides the best fit for their data and make informed decisions based on the results.

Perform a Likelihood Ratio Test in R

A likelihood ratio test compares the goodness of fit of two nested regression models.

A is simply one that contains a subset of the predictor variables in the overall regression model.

For example, suppose we have the following regression model with four predictor variables:

$$Y = \beta_0 + \beta_1x_1 + \beta_2x_2 + \beta_3x_3 + \beta_4x_4 + \varepsilon$$

One example of a nested model would be the following model with only two of the original predictor variables:

$$Y = \beta_0 + \beta_1x_1 + \beta_2x_2 + \varepsilon$$

To determine if these two models are significantly different, we can perform a likelihood ratio test which uses the following null and alternative hypotheses:

H0: The full model and the nested model fit the data equally well. Thus, you should use the nested model.

HA: The full model fits the data significantly better than the nested model. Thus, you should use the full model.

If the p-value of the test is below a certain significance level (e.g. 0.05), then we can reject the null hypothesis and conclude that the full model offers a significantly better fit.

The following example shows how to perform a likelihood ratio test in R.

Example: Likelihood Ratio Test in R

The following code shows how to fit the following two regression models in R using data from the built-in mtcars dataset:

Full model: $\text{mpg} = \beta_0 + \beta_1\text{disp} + \beta_2\text{carb} + \beta_3\text{hp} + \beta_4\text{cyl}$

Reduced model: $\text{mpg} = \beta_0 + \beta_1\text{disp} + \beta_2\text{carb}$

We will use the `lrtest()` function from the `lmtest` package to perform a likelihood ratio test on these two models:

```
library(lmtest)
```

```
#fit full model
```

```
model_full <- lm(mpg ~ disp + carb + hp + cyl, data =  
mtcars)
```

```
#fit reduced model
```

```
model_reduced <- lm(mpg ~ disp + carb, data = mtcars)
```

```
#perform likelihood ratio test for differences in models
```

```
lrtest(model_full, model_reduced)
```

Likelihood ratio test

Model 1: mpg ~ disp + carb + hp + cyl

Model 2: mpg ~ disp + carb

```
#Df LogLik Df Chisq Pr(>Chisq)
```

```
1 6 -77.558
```

```
2 4 -78.603 -2 2.0902 0.3517
```

Since this p-value is not less than .05, we will fail to reject the null hypothesis.

This means the full model and the nested model fit the data equally well. Thus, we should use the nested model because the additional predictor variables in the

full model don't offer a significant improvement in fit.

We could then carry out another likelihood ratio test to determine if a model with only one predictor variable is significantly different from a model with the two predictors:

```
library(lmtest)
```

```
#fit full model
```

```
model_full <- lm(mpg ~ disp + carb, data = mtcars)
```

```
#fit reduced model
```

```
model_reduced <- lm(mpg ~ disp, data = mtcars)
```

```
#perform likelihood ratio test for differences in models
```

```
lrtest(model_full, model_reduced)
```

Likelihood ratio test

```
Model 1: mpg ~ disp + carb
```

```
Model 2: mpg ~ disp
```

```
#Df LogLik Df Chisq Pr(>Chisq)
```

```
1 4 -78.603
```

```
2 3 -82.105 -1 7.0034 0.008136 **
```

```
---
```

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

From the output we can see that the p-value of the likelihood ratio test is 0.008136. Since this is less than .05, we would reject the null hypothesis.

Thus, we would conclude that the model with two predictors offers a significant improvement in fit over the model with just one predictor.

Thus, our final model would be:

$$\text{mpg} = \beta_0 + \beta_1 \text{disp} + \beta_2 \text{carb}$$

How to Perform Simple Linear Regression in R

How to Perform Multiple Linear Regression in R

How to Interpret Significance Codes in R