

Likelihood Ratio Test in R: A Step-by-Step Guide to Model Comparison

Authored by
Mohammed loot

November 6, 2025

RECOMMENDED CITATION

Mohammed loot (2025). *Likelihood Ratio Test in R: A Step-by-Step Guide to Model Comparison*. PSYCHOLOGICAL STATISTICS. Retrieved from <https://statistics.arabpsychology.com/?p=11606>

The [Likelihood Ratio Test](#) (LRT) is a cornerstone of frequentist statistics, providing a robust methodology for comparing the fitness of two statistical [regression models](#). In the complex world of data analysis and predictive modeling, researchers frequently face the challenge of selecting the best model--one that successfully balances explanatory power with essential statistical parsimony. The LRT offers a formal, quantitative method to address this challenge, specifically determining if the added complexity of a larger model is statistically justified compared to a simpler, constrained alternative. This guide provides a detailed walkthrough of how to execute and interpret the LRT effectively within the statistical programming environment, [R](#).

The core principle behind the LRT is based on comparing the maximum likelihood achieved by each model under scrutiny. It quantifies how much the maximized value of the [Maximum Likelihood Estimation](#) (MLE) function decreases when specific predictor variables are removed. Essentially, the test evaluates the evidence against the null hypothesis that the simpler model is just as good as the complex model. By comparing the ratio of these maximized likelihoods, we can ascertain whether the removed parameters contribute meaningfully to the model's ability to explain the observed data variability. This makes the [Likelihood Ratio Test](#) an indispensable tool for rigorous statistical model selection.

The Essential Prerequisite: Understanding Nested Models

For the [Likelihood Ratio Test](#) to be mathematically sound and statistically valid, a crucial prerequisite must be met: the models being compared must be **nested**. Understanding the concept of nesting is paramount before attempting the test. A model is considered to be a [nested model](#) if it can be derived directly from the larger, more comprehensive "full" model simply by setting one or more parameters (coefficients) equal to zero or some other fixed constant.

In the practical application of [regression models](#), the concept of nesting implies that the reduced model must contain a strict subset of the predictor variables present in the full model. If the two models utilize entirely different sets of predictors, or if one model is simply a re-specification of the other (e.g., using a different probability distribution), the LRT cannot be applied. The test statistic relies critically on the difference in degrees of freedom, which must correspond precisely to the number of parameters being constrained or eliminated.

Consider the following general full model designed to predict the dependent variable Y using four predictors (x_1 through x_4), along with an intercept (β_0) and an error term (ϵ):

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

A perfectly valid nested, or reduced, model would be one where we hypothesize that the coefficients for x_3 and x_4 are zero ($\beta_3 = 0$ and $\beta_4 = 0$). This constraint leads

directly to the reduced model:

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

The core inquiry of the LRT is whether forcing those extra parameters (β_3 and β_4) to zero significantly degrades the model's overall fit to the observed data. If the reduction in fit is determined to be negligible, the simpler, more parsimonious model is preferred for its efficiency and ease of interpretation.

Setting Up the Statistical Framework and Hypotheses

To utilize the [Likelihood Ratio Test](#) effectively, we must first establish a clear framework of competing hypotheses that govern our decision-making process. The test is designed to statistically weigh the evidence in favor of the full, more complex model against the simpler, nested model.

H₀ (Null Hypothesis): The null hypothesis posits that the full model and the nested model fit the observed data equally well. This means that the additional parameters included in the full model (those constrained to zero in the reduced model) do not provide a statistically meaningful contribution to the model's explanatory power. If we fail to reject H₀, we conclude that we should **use the nested model** due to the principle of parsimony.

H_A (Alternative Hypothesis): The alternative hypothesis claims that the full model achieves a significantly better fit to the data than the nested model. This suggests that the additional parameters are necessary and provide a meaningful contribution to understanding the relationships within the data. If we reject H₀, we conclude that we should **use the full model**.

The test statistic for the LRT is calculated based on the difference in the log-likelihoods of the two models. Under the null hypothesis, this statistic approximately follows a [Chi-squared distribution](#), where the degrees of freedom are equal to the difference in the number of parameters between the two models. Our decision hinges on the resulting [p-value](#). If the p-value falls below our predefined [significance level](#) (alpha, typically set at $\alpha = 0.05$), we reject the null hypothesis and support the full model. If the p-value exceeds alpha, we fail to reject H₀, favoring the parsimonious nested model.

Practical Application in R: Initial Model Comparison

Executing the Likelihood Ratio Test is a straightforward process when utilizing the statistical programming language [R](#). For this practical demonstration, we will employ the classic built-in dataset, **mtcars**, which compiles various technical specifications and performance metrics for automobiles. Our objective is to determine the most effective combination of predictors for

modeling miles per gallon (mpg).

We begin by defining two competing models that meet the nested requirement. The full model incorporates four potential predictors, while the reduced model contains only two of those predictors. We are specifically testing whether `hp` (horsepower) and `cyl` (number of cylinders) are necessary additions to a model already containing `disp` (engine displacement) and `carb` (number of carburetors).

Full Model (Model 1): `mpg` explained by `disp`, `carb`, `hp`, and `cyl`.

Reduced Model (Model 2): `mpg` explained by only `disp` and `carb`.

To execute the LRT in R, we rely on the `lrtest()` function, which is conveniently packaged within the specialized [lmtest package](#). After loading the necessary library, we fit both models using R's standard `lm()` function for linear regression. Once both models are fitted--ensuring they utilize the same underlying dataset and estimation method--we apply the likelihood ratio test function, placing the full model first, followed by the reduced model, as shown in the code block below.

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
```

Deciphering the LRT Results and Drawing Conclusions

Interpreting the statistical output generated by the `lrtest()` function requires a careful review of several key metrics. These metrics allow us to quantitatively assess the trade-off between model complexity and explanatory power, leading to an evidence-based decision regarding model

selection.

The output table provides comparative statistics for both Model 1 (Full) and Model 2 (Reduced), which are vital for interpretation:

#Df (Degrees of Freedom): This indicates the total number of estimated parameters in the model, including the intercept. For Model 1, the value of 6 corresponds to the intercept plus the four predictors, plus the estimated variance parameter.

LogLik (Maximized Log-Likelihood): This is the maximized value of the log-likelihood function. A higher value (less negative) inherently suggests a better fit for the data, as the goal of MLE is to maximize this function.

Df (Difference in Degrees of Freedom): This crucial column represents the difference in the number of parameters between the full model (Model 1) and the reduced model (Model 2). The value of -2 indicates that two parameters (`hp` and `cyl`) were constrained to zero in the reduced model, and these are the parameters being jointly tested by the LRT.

Chisq: This is the computed value of the [Chi-squared distribution](#) test statistic, derived from the difference between the two log-likelihood values.

Pr(>Chisq) (The p-value): This is the probability associated with the calculated test statistic under the null hypothesis. It is the definitive value used for hypothesis testing.

In the results of our initial comparison, the resulting [p-value](#) is calculated as **0.3517**. Given that this value is substantially larger than the conventional [significance level](#) (α) of 0.05, we must consequently fail to reject the null hypothesis (H_0). The statistical implication of this outcome is that the inclusion of the additional variables (`hp` and `cyl`) in the full model does not yield a statistically significant improvement in explanatory fit compared to the simpler, nested model. Following the principle of statistical parsimony, the analysis suggests the reduced model, $mpg = \beta_0 + \beta_1 disp + \beta_2 carb$, is the optimal choice among these two alternatives.

Iterative Model Refinement and Selection

Model selection is typically an iterative process, not a final single test. Since our previous LRT concluded that the reduced model (Model 2: `mpg ~ disp + carb`) was sufficient and superior to the more complex Model 1, we now treat this successful model as our new "full" reference point and seek to simplify it further. This rigorous approach ensures that we have stripped away all non-contributing variables to arrive at the most efficient predictive structure.

Our next step involves testing whether removing one more predictor, `carb`, results in a significant degradation of fit. We structure our new comparison as follows, focusing on whether `carb` is a necessary predictor:

Full Model (Model 1 - New Baseline): `mpg` explained by `disp` and `carb`.

Reduced Model (Model 2 - Simplified): `mpg` explained by only `disp`.

We repeat the procedure using the `lrtest()` function in R, comparing the fit statistics of the two new models. This iterative approach rigorously validates that only variables that genuinely contribute to predictive accuracy are retained in the final selected model.

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
```

Upon reviewing the output of this second test, we find a markedly different outcome. The resulting [p-value](#) for the likelihood ratio test is calculated as **0.008136**. Since this value is considerably lower than the critical threshold of 0.05, we successfully reject the null hypothesis (H_0). The statistical evidence strongly indicates that the full model (Model 1: `disp + carb`) is a statistically superior fit compared to the model containing only `disp` (Model 2).

The rejection of H_0 confirms that the variable `carb` provides a necessary and significant contribution to the predictive power of the model for `mpg`. Removing it would lead to a substantial and statistically unacceptable degradation of model fit. Therefore, based on the sequential application of the [Likelihood Ratio Test](#), the chosen final model--the one that achieves the best balance between complexity and predictive accuracy--is the two-predictor model:

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

Additional Resources for R Modeling

The [Likelihood Ratio Test](#) stands as a powerful technique for formal model comparison, particularly when dealing with complex datasets where multiple predictor variables are available. By systematically comparing nested models and relying on the resulting Chi-squared statistic, analysts can move beyond subjective judgment and select the most parsimonious model that maintains maximum explanatory power. Mastering this test is essential for anyone engaged in rigorous statistical modeling and regression analysis.

To further advance your proficiency in statistical methods and their implementation within R, we recommend exploring the following resources on linear modeling techniques:

[How to Perform Simple Linear Regression in R](#)

[How to Perform Multiple Linear Regression in R](#)

[How to Interpret Significance Codes in R](#)