

# How to Easily Perform a Wald Test in R

Authored by  
**stats writer**

December 2, 2025

## RECOMMENDED CITATION

stats writer (2025). *How to Easily Perform a Wald Test in R*. PSYCHOLOGICAL SCALES.  
Retrieved from <https://scales.arabpsychology.com/?p=103514>

The Wald Test is a fundamental statistical test indispensable for modern econometric and statistical modeling. Its primary application is to evaluate restrictions on parameters within a model, effectively allowing researchers to determine if certain coefficients or combinations of coefficients are significantly different from hypothesized values, typically zero. While the general concept relates to comparing restricted and unrestricted models, the Wald Test is fundamentally based on comparing the maximum likelihood estimates (MLEs) of the parameters under the unrestricted model. It leverages the asymptotic properties of the maximum likelihood estimator, making it highly versatile across various model types, including linear, generalized linear, and non-linear models. This test is a cornerstone for validating and refining statistical models in the R environment.

The theoretical underpinning of the Wald Test is closely related to both the Likelihood Ratio Test and the Score Test, forming part of the "Holy Trinity" of asymptotic tests in econometrics. The test statistic is calculated by observing how far the unrestricted estimate is from the value specified under the null hypothesis, scaled by the estimated variance of the difference. When dealing with complex models or multiple simultaneous constraints, the resulting statistic follows an approximate chi-squared statistic distribution under the null hypothesis. Understanding how to implement this test effectively in statistical software, such as R, is crucial for rigorous model refinement and validation.

While the standard R function `anova` can perform a general Wald Test for comparing nested linear models (which is equivalent to an F-test for OLS regression), specialized packages and functions are often required for more complex applications, such as generalized linear models (GLMs) or for testing specific subsets of coefficients simultaneously. This guide focuses on the latter, demonstrating how to isolate and test specific subsets of predictors efficiently, a practice essential for achieving parsimony and improving model generalizability in statistical analysis.

## Theoretical Foundations of the Wald Test

The core utility of the **Wald Test** lies in its capacity to rigorously assess whether a linear combination of model parameters differs significantly from zero, or some other specified constant. This flexibility makes it indispensable for testing complex hypotheses beyond simple t-tests for individual coefficients. For instance, one might test if two coefficients are equal ( $\beta_1 = \beta_2$ ) or, more commonly, if a subset of parameters jointly contributes nothing to the predictive power of the model. In essence, the test evaluates the restriction imposed by the null hypothesis against the full, unrestricted model, providing an objective measure of the explanatory value of a group of variables.

The most frequent application, and the focus of this tutorial, involves testing whether a set of predictor variables should be retained in the model. This is achieved by testing if the corresponding regression coefficients are simultaneously equal to zero. If the coefficients are indeed zero, the

variables offer no statistically significant explanatory power relative to the other variables included in the model, and they can be safely omitted. Removing non-significant variables often leads to a more parsimonious and interpretable model structure, which is crucial for maximizing predictive stability and ensuring efficiency in statistical inference.

The test statistic itself is constructed as a quadratic form involving the difference between the estimated parameters and the hypothesized values, weighted by the inverse of the variance-covariance matrix (VCOV) of the estimates. This scaling factor accounts for the uncertainty and correlation among the parameter estimates. This framework is vital for ensuring statistical rigor and preventing Type I or Type II errors during model refinement. We are essentially asking: Does the evidence gathered from our sample suggest that the population parameters violate the restrictions imposed by the null hypothesis, given the inherent sampling variability?

## Formulation of Hypotheses for Parameter Restrictions

When performing a joint hypothesis test using the Wald method, precise articulation of the Null Hypothesis ( $H_0$ ) and the Alternative Hypothesis ( $H_A$ ) is required. For the specific case of parameter subset testing--that is, determining if two or more predictor variables have zero effect on the dependent variable--the hypotheses are stated formally as follows:

**$H_0$ :** The set of predictor variables specified all have regression coefficients equal to zero. For example,  $H_0: \beta_3 = 0$  and  $\beta_4 = 0$ . This implies that the restricted model, which excludes these variables, is statistically adequate.

**$H_A$ :** At least one predictor variable in the specified set has a regression coefficient that is not equal to zero. This suggests that the full, unrestricted model provides a statistically significant improvement over the restricted model, and the subset of variables should be retained.

The statistical decision rule revolves around the calculated Wald test statistic and the associated degrees of freedom, which correspond exactly to the number of restrictions being tested. For a Wald Test evaluating  $q$  restrictions, the degrees of freedom is  $q$ . If the calculated Wald statistic is large, the resulting p-value will be small, suggesting that the observed parameters are far from the hypothesized zero values relative to their sampling variance. This provides strong evidence to reject the Null Hypothesis. Conversely, a small Wald statistic indicates that the data are consistent with the restrictions imposed by  $H_0$ .

Failing to reject the Null Hypothesis carries a significant implication for model building: we are statistically justified in removing the specified set of predictor variables from the model. Their exclusion does not result in a statistically significant degradation of the model's fit. This process is crucial for achieving model parsimony, which favors simpler models that explain the data well, reducing issues like multicollinearity and improving predictive stability, particularly in the context of Multiple Linear Regression.

## Setting Up a Multiple Regression Example in R

To provide a clear, practical demonstration of the Wald Test, we will utilize the robust statistical capabilities of the R environment. Our example will employ the highly accessible built-in dataset, `mtcars`, which compiles data on fuel consumption and ten aspects of automobile design and performance for 32 different models. We aim to construct a model that explains fuel efficiency (miles per gallon, or `mpg`) as a function of four selected vehicle characteristics.

We will fit the following Multiple Linear Regression model, where `$mpg` is the dependent variable:  $\text{mpg} = \beta_0 + \beta_1 \text{disp} + \beta_2 \text{carb} + \beta_3 \text{hp} + \beta_4 \text{cyl}$ . Our specific goal in the subsequent steps is to perform a joint test to see if the variables `carb` (number of carburetors) and `hp` (horsepower) are simultaneously non-significant. This corresponds to testing the joint restriction:  $H_0: \beta_{\text{carb}} = 0 \text{ and } \beta_{\text{hp}} = 0$ . This requires testing the coefficients located at indices 3 and 4 in the full coefficient vector.

## Fitting the Base Model and Interpreting Initial Results

The initial step involves using the standard `lm()` function in R to estimate the parameters of our specified regression model. The model summary provides crucial initial information, including the estimates ( $\hat{\beta}$ ), standard errors, individual t-tests for each coefficient, and overall model fit diagnostics. While the individual t-tests are helpful, they only assess the significance of a single coefficient holding all other predictors constant. They cannot account for the combined explanatory power or the correlation structure when evaluating a subset of variables jointly, which is why the Wald Test is necessary.

The following R code executes the model fitting process and displays the summary statistics. Note that the coefficient vector generated by this model will be ordered as: 1. (Intercept), 2. `disp`, 3. `carb`, 4. `hp`, 5. `cyl`. This ordering dictates the indices we must use later when specifying the terms for the joint Wald Test.

```
#fit regression model
```

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

```
#view model summary
```

```
summary(model)
```

```
Call:
```

```
lm(formula = mpg ~ disp + carb + hp + cyl, data = mtcars)
```

```
Residuals:
```

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

```
-5.0761 -1.5752 -0.2051 1.0745 6.3047
```

Coefficients:

```
Estimate Std. Error t value Pr(>|t|)
(Intercept) 34.021595 2.523397 13.482 1.65e-13 ***
disp -0.026906 0.011309 -2.379 0.0247 *
carb -0.926863 0.578882 -1.601 0.1210
hp 0.009349 0.020701 0.452 0.6551
cyl -1.048523 0.783910 -1.338 0.1922
```

---

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

Residual standard error: 2.973 on 27 degrees of freedom

Multiple R-squared: 0.788, Adjusted R-squared: 0.7566

F-statistic: 25.09 on 4 and 27 DF, p-value: 9.354e-09

Reviewing the individual t-tests, the coefficients for `carb` (p-value = 0.1210) and `hp` (p-value = 0.6551) are both clearly non-significant at the  $\alpha=0.05$  level. This suggests they might be good candidates for removal. However, relying solely on individual t-tests can be misleading due to potential multicollinearity or complex interaction effects. Therefore, employing the joint Wald Test is the statistically rigorous method to confirm that these two variables, when evaluated together, contribute negligible explanatory power to the model.

## Implementing the Wald Test Using the 'aod' Package

To conduct a targeted Wald Test for specific, arbitrary sets of coefficients beyond what the base R installation provides, we rely on specialized packages. We will utilize the `aod` package (Analysis of Overdispersed Data), which offers the versatile `wald.test()` function. This function is perfectly suited for testing joint linear hypotheses on regression coefficients derived from various model classes, including the linear model fitted by `lm()`.

Our objective is to test the joint hypothesis that the regression coefficients for the predictor variables "carb" and "hp" are both equal to zero. The `wald.test()` function requires three primary inputs derived directly from our fitted `model` object: the variance-covariance matrix (VCOV) of the coefficients, the vector of estimated coefficients, and the indices specifying which terms to test. Based on the model summary, the coefficients corresponding to `carb` and `hp` are located at indices 3 and 4, respectively, in the full coefficient vector (following the intercept at index 1 and `disp` at index 2).

## Deconstructing the `wald.test()` Function Syntax

The `wald.test()` function syntax is designed for precision, ensuring the user explicitly defines the components necessary for the test calculation against the asymptotic chi-squared statistic distribution. The fundamental syntax structure is as follows:

### **wald.test(Sigma, b, Terms)**

Understanding each argument is essential for successful and accurate implementation:

**Sigma:** This input represents the estimated variance-covariance matrix of the regression coefficients. This matrix is critical because it captures the correlation between the parameter estimates, allowing the test to correctly scale the discrepancy between the estimated and hypothesized values. In R, we extract this using the `vcov(model)` function.

**b:** This is the vector containing the estimated regression coefficients ( $\hat{\beta}$ ) obtained from the fitted model using `coef(model)`. These are the observed values we are testing against the null hypothesis (zero).

**Terms:** This argument is a numerical vector specifying the indices (positions) of the coefficients that are involved in the joint test. To test the joint significance of `carb` (index 3) and `hp` (index 4), we use the vector `3:4`. The indices must accurately reflect the order of variables in the model output.

The following code demonstrates the practical execution of the Wald Test, specifically testing the joint significance of the coefficients for `carb` and `hp`:

### **library(aod)**

```
#perform Wald Test to determine if the coefficients for 'carb' (3) and 'hp' (4) are both zero
wald.test(Sigma = vcov(model), b = coef(model), Terms = 3:4)
```

Wald test:

-----

Chi-squared test:

$X^2 = 3.6$ ,  $df = 2$ ,  $P(> X^2) = 0.16$

## Interpreting the Chi-Squared Output and Making Decisions

The output generated by the `wald.test()` function provides the necessary metrics for making a statistical decision regarding our joint hypothesis. The output confirms that a Chi-squared test framework was used, providing the Wald test statistic ( $X^2$ ), the degrees of freedom ( $df$ ), and

the corresponding  $P(> X^2)$ , or p-value.

From the results provided, we observe that the calculated Wald test statistic is  $X^2 = 3.6$ , with  $df = 2$ . The degrees of freedom equal 2 because we tested two simultaneous restrictions (that  $\beta_{carb} = 0$  and  $\beta_{hp} = 0$ ). Most importantly, the associated p-value of the test is 0.16. This p-value quantifies the probability of observing a joint difference in coefficients as large as we did, assuming the Null Hypothesis is true.

To conclude the statistical test, we compare the p-value (0.16) to our chosen significance level,  $\alpha$ , typically set at 0.05. Since 0.16 is greater than 0.05, we fail to reject the null hypothesis. The statistical conclusion is clear: we lack sufficient evidence to assert that the coefficients for both `carb` and `hp` are jointly non-zero. This provides strong justification for dropping these terms from the regression model, as their combined contribution does not statistically significantly improve the overall fit relative to a model excluding them.

## Advantages and Limitations of the Wald Test

While the Wald Test is a powerful and frequently used technique in applied statistics, it is essential for the practitioner in R to understand its inherent trade-offs. One major advantage is its computational efficiency: the test only requires fitting the unrestricted model once. All necessary components--the coefficient estimates, their standard errors, and the variance-covariance matrix--are derived directly from this single fit, making it highly efficient for integration into automated model selection routines or large-scale simulation studies.

However, the Wald Test is primarily based on large-sample asymptotic theory. In scenarios where the sample size is limited, or when the underlying distribution of the estimators is highly non-normal, the approximation of the test statistic to the theoretical chi-squared statistic distribution may be inaccurate. This can lead to unreliable p-values and incorrect inferential conclusions, particularly in generalized linear models (GLMs) where the distribution assumptions are more complex than in standard OLS.

Furthermore, a notable limitation of the Wald Test is its sensitivity to model parameterization. It is well-documented that testing equivalent hypotheses expressed under different, but mathematically related, parameterizations can sometimes yield conflicting results--a flaw not shared by the Likelihood Ratio Test. For situations demanding maximum accuracy, such as clinical trials or high-stakes economic forecasting, researchers might prefer the Likelihood Ratio Test, which tends to perform better in finite samples and provides invariance to reparameterization. Nonetheless, the computational convenience and general robustness of the Wald Test secure its position as a default and indispensable tool for routine statistical analysis.