

# Learn How to Test for Heteroscedasticity Using the Goldfeld-Quandt Test in R

Authored by  
**Mohammed looti**

November 6, 2025

## RECOMMENDED CITATION

Mohammed looti (2025). *Learn How to Test for Heteroscedasticity Using the Goldfeld-Quandt Test in R*. PSYCHOLOGICAL STATISTICS. Retrieved from <https://statistics.arabpsychology.com/?p=11546>

## Diagnosing Model Reliability: Heteroscedasticity and the Goldfeld-Quandt Test

One of the fundamental challenges in statistical modeling, particularly when using **Ordinary Least Squares (OLS)** regression, is ensuring the underlying assumptions are met. A critical assumption relates to the variance of the error terms, which must remain constant across all levels of the predictor variables. When this assumption is violated, the condition is known as **heteroscedasticity** (or non-constant variance), a state that severely compromises the validity of subsequent statistical inferences, including parameter estimates and hypothesis testing.

To diagnose this potentially debilitating issue, statisticians rely on specialized diagnostic tools. The **Goldfeld-Quandt test** (GQT) stands out as a powerful and widely accepted method for formally testing the presence of non-constant variance within a linear **regression model**. This test is specifically designed to compare the variance of the **residuals** in different segments of the data, providing a formal statistical measure to determine if the dispersion of errors is indeed uniform across the entire range of observations.

Understanding and addressing **heteroscedasticity** is not merely an academic exercise; it is vital for ensuring the reliability of any predictive or explanatory model. If this variance inconsistency is left uncorrected, the coefficients derived from the model may still be unbiased, but their associated **standard errors** will be skewed and unreliable, leading to potentially incorrect conclusions about the significance of the independent variables. This tutorial provides a detailed, procedural walkthrough of how to execute the **Goldfeld-Quandt test** efficiently using the statistical programming environment **R**, allowing analysts to quickly validate the critical assumption of **homoscedasticity**.

### Preparation in R: Building the Regression Model

Before any diagnostic tests can be performed, a properly specified linear **regression model** must be established. For demonstration purposes, we will utilize the widely available **mtcars** dataset, which is conveniently built into **R**. This dataset contains observations detailing 32 different characteristics of various automobiles, making it an excellent candidate for initial statistical exploration.

Our objective is to model miles per gallon (mpg) as the dependent variable, predicted by engine displacement (disp) and gross horsepower (hp). This specific relationship allows us to analyze how internal engine characteristics influence fuel efficiency. We construct this model using the standard `lm()` function, which is the foundational tool for linear modeling in R. This function computes the Ordinary Least Squares estimates necessary for our analysis.

Executing the `lm()` function and reviewing the subsequent summary output confirms the structural

integrity of the model and provides initial parameter estimates and fit statistics. This preliminary step ensures that the model is correctly fitted to the data before we move on to the more advanced diagnostic phase where we formally test for non-constant variance. The R code snippet below illustrates the fitting process and the resulting coefficient table, which forms the necessary foundation for our subsequent Goldfeld-Quandt analysis.

```
#fit a regression model
model <- lm(mpg~disp+hp, data=mtcars)

#view model summary
summary(model)

Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept) 30.735904 1.331566 23.083 < 2e-16 ***
disp -0.030346 0.007405 -4.098 0.000306 ***
hp -0.024840 0.013385 -1.856 0.073679 .
---
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
Residual standard error: 3.127 on 29 degrees of freedom
Multiple R-squared: 0.7482, Adjusted R-squared: 0.7309
F-statistic: 43.09 on 2 and 29 DF, p-value: 2.062e-09
```

## The Mechanism of the Goldfeld-Quandt Test

The [Goldfeld-Quandt test](#) operates based on a sophisticated methodology designed to isolate potential variance discrepancies within the dataset. The core objective is to divide the total set of observations into two distinct sub-samples--one corresponding to the lower values of the predictor variables and one corresponding to the higher values--and then compare the variance of the errors derived from these two groups.

Crucially, the observations must first be sorted based on the variable suspected of causing [heteroscedasticity](#), which corresponds to the `order.by` argument in R. After sorting, a specified fraction of the observations located centrally in the dataset are intentionally omitted. This removal of the central data points is a necessary step that maximizes the statistical power of the test by creating a greater separation between the variances of the two resulting sub-samples (low range and high range), thereby sharpening the diagnostic capability to detect subtle differences in error dispersion.

Once the two trimmed sub-samples are defined, separate OLS regressions are run on each, and

the residual sum of squares (RSS) is calculated for both. The GQT statistic is then calculated as the ratio of the RSS from the second sub-sample (higher ordered values) to the RSS of the first sub-sample (lower ordered values). This ratio follows an F-distribution under the null hypothesis of constant variance, enabling a formal determination of whether the variances are statistically different enough to reject the assumption of [homoscedasticity](#).

## Execution in R: Utilizing the `lmtest` Package

Implementing the [Goldfeld-Quandt test](#) in [R](#) requires the installation and loading of the highly functional **lmtest** package, which provides the necessary `gqtest()` function. This function efficiently handles the complex process of sorting, splitting, running sub-sample regressions, and calculating the final F-statistic needed for the diagnostic conclusion.

The syntax for the `gqtest()` function requires careful specification of its four primary arguments to ensure the test is executed correctly according to the suspected variance structure:

**model:** This is the object representing the linear [regression model](#) created via `lm()` (our `model` object).

**order.by:** Defines the variable(s) used to sort the data. When multiple predictors are specified (as in our example `~disp+hp`), the test sorts based on the combination of these values, isolating the suspected direction of increasing variance.

**data:** The original dataframe (e.g., **mtcars**) from which the model was derived.

**fraction:** This dictates the number of central observations to be dropped. A standard recommendation is to exclude approximately 20% of the total sample size to achieve optimal test power, ensuring a clear contrast between the low and high variance groups.

Given the **mtcars** dataset has 32 observations, removing 7 observations (which equates to 21.8%) is an appropriate and common choice for the `fraction` parameter. For our specific analysis, we order the data based on both `disp` (displacement) and `hp` (horsepower) simultaneously, providing a comprehensive assessment of the variance structure across increasing levels of engine characteristics.

### #load lmtest library

```
library(lmtest)
```

```
#perform the Goldfeld Quandt test
```

```
gqtest(model, order.by = ~disp+hp, data = mtcars, fraction = 7)
```

```
Goldfeld-Quandt test
```

```
data: model
```

GQ = 1.0316, df1 = 10, df2 = 9, p-value = 0.486

alternative hypothesis: variance increases from segment 1 to 2

## Interpreting the Statistical Output

The resulting output from the `gqtest()` function provides the necessary statistics to formally accept or reject the null hypothesis regarding variance constancy. The key metrics reported are the calculated F-statistic (labeled GQ) and its corresponding p-value. In our example using the `mtcars` data, the resulting test statistic (GQ) is **1.0316**, and the associated p-value is **0.486**.

To accurately interpret these findings, we must reference the formal hypotheses of the [Goldfeld-Quandt test](#). The test is constructed to determine if the desired condition of constant variance is present versus the problematic condition of non-constant variance:

**Null Hypothesis (H0):** The variance of the [residuals](#) is constant across all segments of the data, meaning [homoscedasticity](#) holds true.

**Alternative Hypothesis (HA):** The variance of the residuals is not constant, indicating that [heteroscedasticity](#) is present in the model.

The decision rule dictates that if the p-value is less than the predetermined significance level (typically 0.05), we reject the null hypothesis. Since our calculated p-value of 0.486 is substantially larger than 0.05, we **fail to reject H0**. This outcome strongly suggests that, based on the evidence provided by the GQT, there is insufficient statistical basis to conclude that [heteroscedasticity](#) is a significant issue in our linear [regression model](#). Consequently, we can proceed confidently with the interpretation of the original OLS estimates and their associated [standard errors](#).

## Mitigating Heteroscedasticity: Solutions for Invalid Models

While our specific analysis demonstrated a scenario where the assumption of [homoscedasticity](#) holds, analysts frequently encounter situations where the test mandates the rejection of the null hypothesis (i.e., p-value < 0.05). Confirming the presence of [heteroscedasticity](#) necessitates corrective action, as ignoring this issue renders the calculated [standard errors](#) unreliable, potentially leading to erroneous hypothesis tests and confidence intervals.

Fortunately, statistical theory offers robust methods to address and correct non-constant variance. The choice of mitigation strategy often depends on the severity and underlying cause of the variance issue. Two highly effective approaches involve modifying either the data or the estimation technique itself, ensuring that the final [regression model](#) yields valid and efficient results:

**Transforming the Response Variable.** This involves applying a mathematical transformation to the dependent (response) variable. A frequent and effective technique is taking the logarithm of the

response variable. This log transformation often compresses the scale of the response variable, stabilizing the variance structure of the error terms, thereby resolving the heteroscedasticity issue.

**Utilizing Weighted Least Squares (WLS).** Weighted Least Squares (WLS) regression is an advanced method where each data point is assigned a specific weight based on the estimated variance of its fitted value. Data points associated with higher variance--which are the source of heteroscedasticity--are assigned smaller weights. This effectively reduces the influence of observations with large squared [residuals](#). When appropriate weights are correctly implemented, weighted regression provides efficient parameter estimates and reliable [standard errors](#), eliminating the problem.

For further exploration of statistical diagnostics and advanced techniques for linear modeling, consult the following resources.