

# Learning to Identify Outliers in Linear Regression Models Using the Bonferroni Test in R

Authored by  
**Mohammed looti**

November 13, 2025

## RECOMMENDED CITATION

Mohammed looti (2025). *Learning to Identify Outliers in Linear Regression Models Using the Bonferroni Test in R*. PSYCHOLOGICAL STATISTICS. Retrieved from <https://statistics.arabpsychology.com/?p=24039>

## The Essential Role of Outlier Detection in Regression Analysis

It is fundamentally necessary in the field of statistical modeling to check for [outlier observations](#) when fitting a [linear regression](#) model. Outliers are defined as data points that are significantly distant from the bulk of other observations. Their presence poses a serious threat to model validity because they can drastically skew the calculated coefficients and inflate the standard errors. This distortion inevitably results in an unstable model that fails to accurately capture the true underlying relationship between the predictor variables and the response variable for the majority of the analyzed data.

The influence exerted by even a small number of aberrant points can compromise the overall goodness-of-fit of the model, potentially leading the analyst to draw inaccurate conclusions regarding the statistical significance of the chosen predictors. Furthermore, if the primary objective of the model is forecasting, these influential outliers can cause substantial problems, resulting in unreliable predictions for the response values of new, unseen observations. Consequently, establishing a systematic and statistically robust approach to identifying and managing outliers is not just recommended, but is a fundamental requirement for sound data analysis and reliable model validation.

Identifying these influential points represents more than a simple exercise in data removal; it demands a critical understanding of their origin. Outliers may signal simple measurement errors, mistakes during data entry, or they might represent genuinely rare, yet naturally occurring, extreme events. If they are clerical errors, they should be corrected or removed. If they represent natural variability, the analyst must determine if the chosen modeling technique is robust enough or if an alternative, less sensitive regression method is warranted. The technique we explore here, the **Bonferroni outlier test**, provides a statistically rigorous framework specifically designed for flagging these suspicious points within the context of a fitted regression model.

## Deconstructing the Bonferroni Outlier Test Methodology

A widely accepted and statistically robust method for checking for outliers in the context of a regression model is the utilization of the **Bonferroni outlier test**. This test operates by critically analyzing the [Studentized residuals](#) (also known as externally studentized residuals) calculated for every observation in the dataset. A [Studentized residual](#) is essentially the standardized residual of an observation, where the standardization is achieved by dividing the residual by its estimated standard deviation, a calculation performed after temporarily removing that specific observation from the dataset. This procedure effectively highlights how extreme an observation is relative to the model fit by all the remaining data points.

The definitive feature of this testing procedure is its incorporation of the [Bonferroni correction](#). When an analyst simultaneously tests multiple data points for outlier status, the inherent probability

of obtaining a false positive (a Type I error) increases substantially--a phenomenon known as the inflation of the family-wise error rate. To counteract this inflation, the [Bonferroni correction](#) adjusts the individual [p-values](#) by multiplying them by the total number of observations (which equals the number of tests performed).

This adjustment is critical because it ensures that the overall probability of incorrectly identifying at least one observation as an outlier remains below the analyst's specified significance level (alpha, conventionally set at 0.05). The test outputs these highly conservative, adjusted **p-values** for each observation. By applying the **Bonferroni correction**, we gain a highly reliable measure that clearly indicates which observations, if any, are true outliers warranting detailed follow-up investigation. If the adjusted p-value falls below the chosen significance threshold (e.g., 0.05), the observation is confidently flagged as a statistically significant outlier.

## Implementing the Test in R: Syntax and Prerequisites

The most straightforward and widely adopted method for performing the **Bonferroni outlier test** within the R environment is through the use of the powerful `outlierTest()` function. This function is conveniently housed within the widely utilized **car package** (Companion to Applied Regression). Prior to executing the test, it is mandatory to ensure that the **car package** has been both installed on your system and loaded into the current R session using the `library()` command.

The `outlierTest()` function is specifically designed to interact with model objects generated using R's standard `lm()` function for linear regression. It requires the previously fitted model object as its primary, mandatory argument. The general syntax structure is remarkably concise and flexible, enabling seamless integration into virtually any existing regression analysis workflow without extensive modification.

The `outlierTest()` function uses the following syntax:

```
outlierTest(model, cutoff=.05, ...)
```

Where the parameters are defined as:

**model:** This is the required argument, representing a [linear regression](#) model object that has been fit using R's standard `lm()` function.

**cutoff:** This optional parameter specifies the statistical significance threshold for the Bonferroni-adjusted p-values. By default, this value is set to **.05**. Observations whose Bonferroni p-values exceed this threshold are generally not explicitly reported in the output, unless the function finds no statistically significant outliers, in which case it reports the single observation exhibiting the largest absolute **Studentized residual**.

It is important to recognize that while the default **cutoff** value of **.05** aligns perfectly with

standard statistical practice and is highly recommended, analysts retain the flexibility to adjust this value. If a specific research context demands a stricter or looser requirement for defining a statistically significant outlier, the threshold can be altered. However, analysts must always exercise caution and provide clear justification when deviating from the conventional 0.05 alpha level.

## Practical Application: Modeling with the mtcars Dataset

To effectively demonstrate the practical application of the `outlierTest()` function, we will utilize a widely known, built-in dataset in R: `mtcars`. This dataset provides comprehensive information covering various performance attributes and physical characteristics for 32 different automobile models from the 1970s. This example meticulously illustrates the standard process of fitting a regression model and subsequently subjecting it to rigorous outlier diagnostics.

Before fitting any model, preliminary inspection of the data structure is a prudent step to ensure familiarity with the variables. We use the `head()` function to view the initial rows of the dataset, confirming variable names (like `mpg`, `disp`, and `carb`) and their typical values:

```
#view head of mtcars dataset
```

```
head(mtcars)
```

```
mpg cyl disp hp drat wt  qsec vs am gear carb
Mazda RX4 21.0 6 160 110 3.90 2.620 16.46 0 1 4 4
Mazda RX4 Wag 21.0 6 160 110 3.90 2.875 17.02 0 1 4 4
Datsun 710 22.8 4 108 93 3.85 2.320 18.61 1 1 4 1
Hornet 4 Drive 21.4 6 258 110 3.08 3.215 19.44 1 0 3 1
Hornet Sportabout 18.7 8 360 175 3.15 3.440 17.02 0 0 3 2
Valiant 18.1 6 225 105 2.76 3.460 20.22 1 0 3 1
```

For the purpose of this diagnostic demonstration, we construct a [multiple linear regression](#) model where we aim to predict `mpg` (miles per gallon) using `disp` (engine displacement) and `carb` (number of carburetors) as the key predictor variables. Constructing and fitting this model is the essential precursor step before initiating any sophisticated outlier analysis. We use the standard R syntax below to fit the model using `lm()` and subsequently review the summary statistics, which provide insight into the model's overall fit and the initial significance of the predictors:

```
#fit first regression model
```

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

```
#view model summary
```

```
summary(fit)
```

Call:

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

Residuals:

Min 1Q Median 3Q Max

```
-4.3379 -2.0849 -0.3448 1.5118 6.2836
```

Coefficients:

Estimate Std. Error t value Pr(>|t|)

```
(Intercept) 31.152710 1.263620 24.654 < 2e-16 ***
```

```
disp -0.036296 0.004676 -7.762 1.47e-08 ***
```

```
carb -0.955677 0.358789 -2.664 0.0125 *
```

---

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

Residual standard error: 2.964 on 29 degrees of freedom

Multiple R-squared: 0.7737, Adjusted R-squared: 0.7581

F-statistic: 49.58 on 2 and 29 DF, p-value: 4.393e-10

With our **linear regression** model successfully fitted and its basic properties reviewed, we can now confidently proceed to the specific task of outlier detection. This diagnostic step is crucial for upholding the necessary assumption of normally distributed residuals and for ensuring that the model's strong metrics (such as the high R-squared value) are not misleadingly inflated or distorted by the presence of a few influential points.

## Executing the Bonferroni Test and Analyzing R Output

Our primary objective is to execute the **Bonferroni outlier test** to systematically determine whether any of the 32 observations in the dataset are considered statistically significant outliers within the context of our established regression model. This requires loading the necessary package and initiating the core diagnostic function.

We begin by loading the **car package** using the `library()` function. Once the package is active, we apply the `outlierTest()` function directly to our fitted model object, named `fit`. This single command initiates the calculation of the [Studentized residuals](#) and the necessary [Bonferroni correction](#) for every data point.

We use the following syntax to perform the diagnostic test:

```
library(car)
```

```
#perform Bonferroni outlier test
outlierTest(fit)
No Studentized residuals with Bonferroni p < 0.05
Largest |rstudent|:
rstudent unadjusted p-value Bonferroni p
Toyota Corolla 2.411735 0.022681 0.72579
```

The output provided by R offers immediate and valuable insight. The very first line summarizes the key finding based on the default 0.05 cutoff level: "**No Studentized residuals with Bonferroni p < 0.05**". This definitive statement confirms that, even after applying the stringent adjustment for multiple comparisons inherent in testing all 32 data points simultaneously, none of the observations meet the strict statistical criteria required to be classified as a significant outlier at the 5% alpha level.

## Interpreting Diagnostic Metrics: The Largest Residual

The immediate and most important conclusion drawn from this test output is the absence of statistically significant outliers influencing this particular **linear regression** model. This favorable finding suggests that the model is relatively stable and that the observed residuals are likely attributable to random error inherent in the data collection rather than highly influential, aberrant data points. Consequently, this outcome significantly increases confidence in the model's reliability for both interpretation and prediction purposes.

Crucially, even when no significant outliers are detected, the `outlierTest()` function provides essential diagnostic information by highlighting the observation that possesses the largest absolute [Studentized residual](#). In our example, this most divergent point is labeled "Toyota Corolla."

The detailed output for this largest residual includes three key values that guide the analyst's interpretation:

**rstudent:** This value (2.411735) represents the externally [Studentized residual](#). This metric measures how many standard deviations the observation is situated away from the regression line predicted by a model fitted exclusively without that specific observation. Higher absolute values strongly suggest a greater degree of divergence.

**unadjusted p-value:** This is the p-value (0.022681) associated with testing whether this single observation is an outlier, calculated without applying any multiple comparison correction. Had the analysis focused solely on this one data point, it would technically be deemed statistically significant at the 0.05 level.

**Bonferroni p:** This is the final, adjusted [p-value](#) (0.72579). Since the unadjusted p-value is multiplied by 32 (or a similar conservative calculation) to control the family-wise error rate, the

resulting Bonferroni p-value is substantially higher. Because 0.72579 is vastly greater than the cutoff of 0.05, we confidently conclude that the "Toyota Corolla," despite being the most extreme point, is not statistically significant as an outlier when controlling for the overall testing error rate via the **Bonferroni correction**.

In summary, the **Bonferroni outlier test** provides an essential layer of diagnostic rigor, ensuring the integrity of a regression analysis. It furnishes analysts with the statistical confidence needed to flag points that genuinely require further investigation, thereby strengthening the validity and generalizability of the resulting statistical model.

## Expanding Diagnostics: Influence and Leverage

While the **Bonferroni outlier test** is highly effective for identifying vertical outliers based primarily on the response variable, statistical analysts must be aware of other critical diagnostics that assess influence and leverage in the design space. A complete and robust analysis typically requires reviewing multiple influence statistics to ensure the model's robustness against all forms of unusual data points. Key metrics in this area include [Cook's Distance](#), which assesses overall influence, and [DFFITS](#), which measures how much the predicted value changes when an observation is removed.

The **car package**, which houses the `outlierTest()` function, provides a comprehensive suite of related functions that can be utilized to perform these other common diagnostic tasks in R. This includes various specialized plotting functions that allow for visual assessment of residuals, leverage, and influence, offering tools for regression assumption checking far beyond the scope of this specific outlier test.

<!--

## Featured Posts

-->