

# Learning Guide: Testing for Autocorrelation in Regression Models Using the Breusch-Godfrey Test with R

Authored by  
**Mohammed looti**

November 5, 2025

## RECOMMENDED CITATION

Mohammed looti (2025). *Learning Guide: Testing for Autocorrelation in Regression Models Using the Breusch-Godfrey Test with R*. PSYCHOLOGICAL STATISTICS. Retrieved from <https://statistics.arabpsychology.com/?p=10236>

## The Critical Assumption of Independent Residuals in OLS Modeling

A cornerstone of classical regression analysis, particularly when utilizing [Ordinary Least Squares \(OLS\)](#), is the assumption that the error terms (or residuals) derived from the model are independently and identically distributed. This independence is not merely a theoretical nicety; it requires that the error associated with one observation must bear no correlation whatsoever with the error associated with any other observation in the dataset.

The fulfillment of this independence criterion is what guarantees that the OLS estimators retain the highly desirable property of being the Best Linear Unbiased Estimators (BLUE). Conversely, when residuals exhibit systematic dependencies or patterns, the calculated standard errors of the coefficient estimates become unreliable. This bias subsequently invalidates critical hypothesis testing procedures and can lead researchers to draw incorrect statistical inference regarding the true significance of the chosen predictor variables.

This critical assumption is frequently violated in practice, especially when analyzing sequential data such as [time series](#) or panel data collected over time. This violation is formally termed serial correlation or [autocorrelation](#), a phenomenon where the error term generated at a given point in time is systematically related to the error terms from previous periods. Detecting and addressing this issue is essential for sound econometric modeling.

### Detecting Autocorrelation Beyond the First Order

Autocorrelation signifies that the residuals are correlated with their own lagged values. This situation is endemic in many financial, economic, and engineering models where sequential observations naturally carry over information from preceding data points. Failing to account for this dependency results in models that dramatically overstate the precision of their estimates, leading to overly optimistic confidence intervals and potentially misleading conclusions about variable significance.

Traditionally, researchers have relied on tests like the [Durbin-Watson test](#) to diagnose serial correlation. However, the Durbin-Watson test is inherently limited, as it is primarily designed to detect simple, first-order autocorrelation--meaning the error is only related to the error immediately preceding it (lag 1).

In many real-world scenarios, particularly with high-frequency or seasonal data, correlation may manifest at higher orders. For instance, quarterly economic data might exhibit a seasonal pattern where errors from the current quarter are correlated with those from the same quarter last year (lag 4). When testing for these complex, higher-order dependencies, a more comprehensive and flexible diagnostic tool is mandatory. This is precisely where the **Breusch-Godfrey test** proves its value, offering the capability to examine autocorrelation up to any specified order, denoted as  $p$ .

## Theoretical Foundation of the Breusch-Godfrey Test

The Breusch-Godfrey (BG) test, often referred to as the Lagrange Multiplier (LM) test for serial correlation, stands as a versatile and robust diagnostic procedure. A key advantage of the BG test over older methods is its applicability even when the regression model includes lagged dependent variables, making it superior for dynamic modeling contexts.

The methodology of the BG test relies on constructing an auxiliary regression. First, the residuals are obtained from the original OLS model. These residuals are then regressed on all the original predictor variables and, crucially, the set of lagged residuals up to the maximum order  $p$  under consideration. The test then assesses whether the coefficients corresponding to these lagged residuals are jointly zero.

The formal hypotheses governing the Breusch-Godfrey test are structured as follows:

**H<sub>0</sub> (Null Hypothesis):** There is no significant [autocorrelation](#) among the residuals for any lag up to the order  $p$ . Statistically, this means all coefficients associated with the lagged residuals in the auxiliary regression are zero.

**H<sub>A</sub> (Alternative Hypothesis):** Significant autocorrelation exists among the residuals at at least one lag less than or equal to  $p$ . This implies that at least one lagged residual coefficient is statistically non-zero.

The test statistic derived from this auxiliary regression approximates a [Chi-Square distribution](#) with  $p$  degrees of freedom. The final decision to reject or fail to reject the null hypothesis is based on comparing the calculated [p-value](#) to a predefined level of significance, commonly set at  $\alpha = 0.05$ .

## Practical Implementation in R using the `lmtest` Package

To perform the Breusch-Godfrey test efficiently within the R statistical environment, we rely on the powerful functions provided by the **lmtest** package. This package is specifically designed for rigorous diagnostic checks on linear models, offering a streamlined interface for executing tests like the BG test.

The central function utilized is `bgtest()`. This function accepts the existing model formula and the desired maximum order of autocorrelation ( $p$ ) as its primary parameters. The standardized syntax structure is typically: **`bgtest(formula, order = p, data = dataset_name)`**. A prerequisite for running this diagnostic is the prior definition and fitting of the linear regression model whose residuals are to be examined.

For demonstration purposes, we first establish a small synthetic dataset suitable for multiple linear

regression. This dataset includes a response variable ( $y$ ) and two predictor variables ( $x_1$  and  $x_2$ ), allowing us to proceed with a standard OLS estimation before applying the diagnostic test.

#### #create dataset

```
df <- data.frame(x1=c(3, 4, 4, 5, 8, 9, 11, 13, 14, 16, 17, 20),
x2=c(7, 7, 8, 8, 12, 4, 5, 15, 9, 17, 19, 19),
y=c(24, 25, 25, 27, 29, 31, 34, 34, 39, 30, 40, 49))
```

```
#view first six rows of dataset
```

```
head(df)
```

```
x1 x2 y
1 3 7 24
2 4 7 25
3 4 8 25
4 5 8 27
5 8 12 29
6 9 4 31
```

## Executing the Test and Interpreting the Results

With the dataset prepared, the next step involves running the Breusch-Godfrey test. In this illustrative scenario, we hypothesize that serial correlation might extend up to the third order ( $p = 3$ ). We must first load the necessary library, `lmtest`, and then execute the `bgtest()` function, ensuring the regression formula  $y \sim x_1 + x_2$  and the order argument are correctly specified.

#### #load lmtest package

```
library(lmtest)
```

```
#perform Breusch-Godfrey test
```

```
bgtest(y ~ x1 + x2, order=3, data=df)
```

```
Breusch-Godfrey test for serial correlation of order up to 3
```

```
data: y ~ x1 + x2
```

```
LM test = 8.7031, df = 3, p-value = 0.03351
```

The resulting output provides all the necessary statistics for conducting the formal hypothesis test. We observe that the Lagrange Multiplier (LM) test statistic, which is asymptotically distributed as a Chi-Square variable, is calculated as **8.7031**, corresponding to 3 degrees of freedom ( $df = p = 3$ ). Most critically, the corresponding [p-value](#) is reported as **0.03351**.

To finalize the conclusion, we compare this p-value to our standard significance threshold,  $\alpha = 0.05$ . Since 0.03351 is demonstrably less than 0.05, we have sufficient statistical evidence to reject the null hypothesis ( $H_0$ ). This definitive rejection leads us to the conclusion that significant serial correlation exists among the residuals at an order less than or equal to 3. This result is a strong warning that the original OLS model estimation is likely inefficient and its standard errors are biased, thereby requiring immediate corrective measures.

## Strategies for Addressing and Correcting Serial Correlation

When the Breusch-Godfrey test confirms the presence of serial correlation, relying on standard OLS output is ill-advised. While the coefficient estimates themselves remain unbiased (provided the correlation is in the error term and not the predictors), the calculated standard errors are systematically incorrect, rendering subsequent t-statistics and F-statistics invalid. Correcting this issue is paramount for achieving a robust and scientifically reliable model.

The appropriate correction strategy often depends on diagnosing the root cause of the correlation. Sometimes, the dependency arises from model misspecification--perhaps a crucial lagged variable was omitted, or seasonality was neglected. In such cases, the primary fix is refining the structural form of the regression equation.

If the model specification appears sound, alternative estimation methodologies or adjustments to the inference procedures must be implemented. A rigorous solution involves employing [Generalized Least Squares \(GLS\)](#), which explicitly models and accounts for the detected correlation structure within the error terms. Alternatively, a simpler and often effective approach is to use [HAC \(Heteroskedasticity and Autocorrelation Consistent\)](#) standard errors, such as the Newey-West estimator, which adjusts the standard errors to provide valid inference without altering the original OLS coefficients.

Specific corrective strategies tailored to the type of residual dependency include:

For **positive serial correlation**, which suggests inertia in the time series, the model should be augmented by including lagged values of the dependent variable and/or specific independent variables to capture that momentum.

For **negative serial correlation**, researchers must carefully scrutinize how variables have been transformed. Negative correlation often signals that the data has been [overdifferenced](#), meaning differencing has been applied too many times, artificially creating an alternating pattern in the errors.

For **seasonal correlation** (common at lag 4 for quarterly data or lag 12 for monthly data), incorporate seasonal dummy variables or seasonal lags of the dependent variable directly into the model structure to absorb the seasonal effect.

As a robust backup, employ generalized estimation methods like GLS, or utilize inference based

on HCC/HAC standard errors (e.g., using the `coefTest` function with `vcov = NeweyWest` in R) to ensure that statistical inference remains valid despite the detected dependency in the residuals.

## Conclusion: Ensuring Robust Inference

The Breusch-Godfrey test represents an indispensable component of the regression diagnostics toolkit, especially when analyzing sequential or [time series](#) data where serial dependence is a high risk. By allowing analysts to test rigorously for autocorrelation at multiple, higher orders, the BG test ensures that the crucial OLS assumption of independent residuals is either validated or flagrantly exposed as violated.

The mandate for practitioners is clear: ensuring that residuals are well-behaved is fundamental to developing models that are not only accurate in prediction but also reliable for drawing statistical inference. A significant result from the **Breusch-Godfrey test** serves as a critical signal, demanding immediate action to either refine the core model specification or transition to specialized estimation techniques that are capable of accommodating the complexity of the detected serial correlation structure.