

# How to Calculate the Durbin-Watson Statistic in Excel for Autocorrelation Testing

Authored by  
**stats writer**

December 6, 2025

## RECOMMENDED CITATION

stats writer (2025). *How to Calculate the Durbin-Watson Statistic in Excel for Autocorrelation Testing*. PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=106249>

The Durbin-Watson test is a critical inferential statistical procedure used primarily to check for the presence of autocorrelation (or serial correlation) within the residuals derived from a time-series regression analysis. Failing to account for this correlation can lead to severely biased standard errors and invalid statistical inferences. To effectively apply this test within Microsoft Excel, the analyst must first successfully execute the core regression model. Following the regression, we utilize a specialized function--or manual calculation methods--to determine the Durbin-Watson statistic, typically denoted as 'd', using the generated residuals. This statistic is bound between 0 and 4. Values that approach 2 strongly suggest a lack of significant autocorrelation, which is the desired outcome. Conversely, values deviating significantly from 2 (near 0 or near 4) provide evidence for positive or negative serial correlation, respectively, requiring further investigation and potential model adjustment.

The integrity and reliability of any statistical model, particularly those based on Ordinary Least Squares regression analysis, hinge upon satisfying several foundational assumptions. One of the most important of these assumptions, especially when dealing with time-series or sequential data, is that the model errors (or residuals) are independent of one another.

A rigorous way to formally assess whether this assumption of independent residuals has been met is by performing the Durbin-Watson test, which is specifically used to detect the presence of first-order autocorrelation in the residuals of a regression analysis. This test uses the following hypotheses to guide the decision process:

**H<sub>0</sub> (The Null Hypothesis):** There is no statistically significant correlation among the residuals (i.e., the errors are independent).

**H<sub>A</sub> (The Alternative Hypothesis):** The residuals are significantly autocorrelated (either positively or negatively).

This tutorial provides a systematic, step-by-step example detailing how to perform a Durbin-Watson test in Excel, from data preparation to final interpretation.

## Step 1: Preparing and Entering the Data for Analysis

Before any statistical calculation can commence, the data must be correctly formatted within the Excel spreadsheet. For a standard multiple linear regression setup, data points should be arranged in adjacent columns, with one column designated as the dependent (response) variable (Y) and the remaining columns serving as the independent (predictor) variables (X). This structure is essential for the seamless execution of the regression procedure using Excel's built-in tools.

In this instructional example, we will utilize a small dataset containing 13 observations. We aim to construct a multiple linear regression model where 'y' is the variable we are trying to predict, based

on the values of the two predictor variables, 'x1' and 'x2'. Ensuring the data is contiguous and clearly labeled will prevent errors during the input phase of the regression tool.

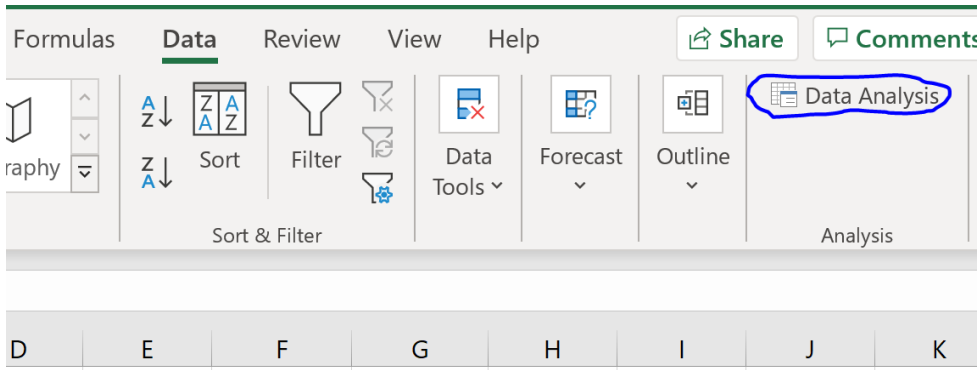
Observe the structure of our sample data below. Note that the dataset includes an index (Observation Number) for clarity, although this column is not used in the regression calculation itself. The variables of interest (Y, X1, X2) are clearly defined:

|    | A         | B         | C        | D | E | F | G |
|----|-----------|-----------|----------|---|---|---|---|
| 1  | <b>x1</b> | <b>x2</b> | <b>y</b> |   |   |   |   |
| 2  | 3         | 4         | 45       |   |   |   |   |
| 3  | 5         | 4         | 47       |   |   |   |   |
| 4  | 5         | 5         | 49       |   |   |   |   |
| 5  | 6         | 5         | 54       |   |   |   |   |
| 6  | 7         | 6         | 50       |   |   |   |   |
| 7  | 8         | 6         | 47       |   |   |   |   |
| 8  | 8         | 7         | 50       |   |   |   |   |
| 9  | 12        | 6         | 59       |   |   |   |   |
| 10 | 14        | 5         | 64       |   |   |   |   |
| 11 | 15        | 6         | 60       |   |   |   |   |
| 12 | 16        | 6         | 58       |   |   |   |   |
| 13 | 16        | 8         | 59       |   |   |   |   |
| 14 | 19        | 7         | 52       |   |   |   |   |
| 15 |           |           |          |   |   |   |   |
| 16 |           |           |          |   |   |   |   |
| 17 |           |           |          |   |   |   |   |
| 18 |           |           |          |   |   |   |   |
| 19 |           |           |          |   |   |   |   |
| 20 |           |           |          |   |   |   |   |
| 21 |           |           |          |   |   |   |   |
| 22 |           |           |          |   |   |   |   |

## Step 2: Fitting a Multiple Linear Regression Model and Extracting Residuals

The next critical phase is fitting the multiple linear regression model. This step is necessary not only for establishing the relationships between the variables but, crucially for the Durbin-Watson test, for generating the model's residuals. These residuals represent the difference between the actual observed values and the values predicted by the model ( $\hat{Y}$ ), and they are the essential input required for the Durbin-Watson calculation.

To initiate the regression process, navigate to the **Data** tab along the top ribbon. Then locate and click **Data Analysis** within the **Analyze** group.



If you don't see this option readily available, you need to first **install and activate the Analysis ToolPak add-in**, as it provides the core statistical functions necessary for the regression.

In the dialogue window that appears, scroll down and click **Regression** and then click **OK**. In the new window that appears, fill in the necessary information, ensuring that your Input Y Range contains the dependent variable and the Input X Range contains all predictor variables. Most importantly, ensure the checkbox for **Residuals** is checked, as we need this output for the next step:

|    | A  | B  | C  | D | E | F | G | H | I |
|----|----|----|----|---|---|---|---|---|---|
| 1  | x1 | x2 | y  |   |   |   |   |   |   |
| 2  | 3  | 4  | 45 |   |   |   |   |   |   |
| 3  | 5  | 4  | 47 |   |   |   |   |   |   |
| 4  | 5  | 5  | 49 |   |   |   |   |   |   |
| 5  | 6  | 5  | 54 |   |   |   |   |   |   |
| 6  | 7  | 6  | 50 |   |   |   |   |   |   |
| 7  | 8  | 6  | 47 |   |   |   |   |   |   |
| 8  | 8  | 7  | 50 |   |   |   |   |   |   |
| 9  | 12 | 6  | 59 |   |   |   |   |   |   |
| 10 | 14 | 5  | 64 |   |   |   |   |   |   |
| 11 | 15 | 6  | 60 |   |   |   |   |   |   |
| 12 | 16 | 6  | 58 |   |   |   |   |   |   |
| 13 | 16 | 8  | 59 |   |   |   |   |   |   |
| 14 | 19 | 7  | 52 |   |   |   |   |   |   |

**Regression** [?] [X]

**Input**

Input Y Range:  [↑]

Input X Range:  [↑]

Labels  Constant is Zero

Confidence Level:  %

**Output options**

Output Range:  [↑]

New Worksheet Ply:

New Workbook

**Residuals**

Residuals  Residual Plots

Standardized Residuals  Line Fit Plots

**Normal Probability**

Normal Probability Plots

[OK] [Cancel] [Help]

Once you click **OK**, the comprehensive regression output will appear, typically on a new worksheet, which includes the coefficients, summary statistics, and a list of the calculated residuals:

| D | E                            | F                   | G                   | H             | I              | J                    | K                | L                 | M                 |
|---|------------------------------|---------------------|---------------------|---------------|----------------|----------------------|------------------|-------------------|-------------------|
|   | SUMMARY OUTPUT               |                     |                     |               |                |                      |                  |                   |                   |
|   | <i>Regression Statistics</i> |                     |                     |               |                |                      |                  |                   |                   |
|   | Multiple R                   | 0.750703            |                     |               |                |                      |                  |                   |                   |
|   | R Square                     | 0.563556            |                     |               |                |                      |                  |                   |                   |
|   | Adjusted R                   | 0.476267            |                     |               |                |                      |                  |                   |                   |
|   | Standard E                   | 4.367605            |                     |               |                |                      |                  |                   |                   |
|   | Observatio                   | 13                  |                     |               |                |                      |                  |                   |                   |
|   | ANOVA                        |                     |                     |               |                |                      |                  |                   |                   |
|   |                              | <i>df</i>           | <i>SS</i>           | <i>MS</i>     | <i>F</i>       | <i>ignificance F</i> |                  |                   |                   |
|   | Regression                   | 2                   | 246.3172            | 123.1586      | 6.456215       | 0.015836             |                  |                   |                   |
|   | Residual                     | 10                  | 190.7597            | 19.07597      |                |                      |                  |                   |                   |
|   | Total                        | 12                  | 437.0769            |               |                |                      |                  |                   |                   |
|   |                              | <i>Coefficients</i> | <i>andard Error</i> | <i>t Stat</i> | <i>P-value</i> | <i>Lower 95%</i>     | <i>Upper 95%</i> | <i>ower 95.0%</i> | <i>pper 95.0%</i> |
|   | Intercept                    | 50.30152            | 6.748134            | 7.454138      | 2.17E-05       | 35.26574             | 65.3373          | 35.26574          | 65.3373           |
|   | x1                           | 1.032967            | 0.323748            | 3.190647      | 0.009644       | 0.311611             | 1.754323         | 0.311611          | 1.754323          |
|   | x2                           | -1.31117            | 1.45364             | -0.90199      | 0.388272       | -4.55008             | 1.927748         | -4.55008          | 1.927748          |
|   | RESIDUAL OUTPUT              |                     |                     |               |                |                      |                  |                   |                   |
|   | <i>Observation</i>           | <i>Predicted y</i>  | <i>Residuals</i>    |               |                |                      |                  |                   |                   |
|   | 1                            | 48.15576            | -3.15576            |               |                |                      |                  |                   |                   |
|   | 2                            | 50.2217             | -3.2217             |               |                |                      |                  |                   |                   |
|   | 3                            | 48.91053            | 0.089467            |               |                |                      |                  |                   |                   |
|   | 4                            | 49.9435             | 4.0565              |               |                |                      |                  |                   |                   |
|   | 5                            | 49.6653             | 0.334698            |               |                |                      |                  |                   |                   |
|   | 6                            | 50.69827            | -3.69827            |               |                |                      |                  |                   |                   |
|   | 7                            | 49.3871             | 0.612897            |               |                |                      |                  |                   |                   |
|   | 8                            | 54.83014            | 4.169863            |               |                |                      |                  |                   |                   |
|   | 9                            | 58.20724            | 5.792764            |               |                |                      |                  |                   |                   |
|   | 10                           | 57.92904            | 2.070962            |               |                |                      |                  |                   |                   |
|   | 11                           | 58.962              | -0.962              |               |                |                      |                  |                   |                   |
|   | 12                           | 56.33967            | 2.660326            |               |                |                      |                  |                   |                   |
|   | 13                           | 60.74974            | -8.74974            |               |                |                      |                  |                   |                   |

### Step 3: Calculating the Durbin-Watson Statistic

The test statistic for the Durbin-Watson test, denoted  $d$ , measures the degree of linear autocorrelation between adjacent errors. It is fundamentally derived by comparing the squared differences between consecutive residuals to the sum of the squared residuals themselves. The formula is calculated as follows:

$$d = \frac{\sum_{t=2}^T (e_t - e_{t-1})^2}{\sum_{t=1}^T e_t^2}$$

where:

**T**: The total number of observations in the dataset (N).

**et**: The  $t^{\text{th}}$  residual from the regression model.

To calculate this test statistic in Excel efficiently, we can use a specialized formula that leverages array calculations over the column of residuals generated in Step 2. Assuming your residuals start at row 2 and continue through row T, the formula below calculates the sum of the squared differences between sequential residuals and divides it by the sum of the squared residuals:

| RESIDUAL OUTPUT |             |           |   |          |   |
|-----------------|-------------|-----------|---|----------|---|
| Observation     | Predicted y | Residuals | d | 1.347536 | =SUMXMY2(G27:G38, G26:G37) / SUMSQ(G26:G38) |
| 1               | 48.15576    | -3.15576  |   |          |   |
| 2               | 50.2217     | -3.2217   |   |          |   |
| 3               | 48.91053    | 0.089467  |   |          |   |
| 4               | 49.9435     | 4.0565    |   |          |   |
| 5               | 49.6653     | 0.334698  |   |          |   |
| 6               | 50.69827    | -3.69827  |   |          |   |
| 7               | 49.3871     | 0.612897  |   |          |   |
| 8               | 54.83014    | 4.169863  |   |          |   |
| 9               | 58.20724    | 5.792764  |   |          |   |
| 10              | 57.92904    | 2.070962  |   |          |   |
| 11              | 58.962      | -0.962    |   |          |   |
| 12              | 56.33967    | 2.660326  |   |          |   |
| 13              | 60.74974    | -8.74974  |   |          |   |

When executed against our sample data, the resulting test statistic turns out to be **1.3475**. This value must now be interpreted against the standard critical values.

## Interpreting the Durbin-Watson Statistic using Critical Values

To determine if a Durbin-Watson test statistic is statistically significant at a certain alpha level, we must refer to a table of critical values ( $d_L$  and  $d_U$ ). These tables define three regions: a region of positive autocorrelation, a region of no autocorrelation, and an inconclusive region between them.

For our analysis, we assume a standard significance level of  $\alpha = .05$ . Given that we have  $n = 13$  observations and  $k = 2$  independent variables in the regression model, consulting the appropriate Durbin-Watson table shows the following critical values for testing positive autocorrelation:

Lower critical value ( $d_L$ ): **0.86**

Upper critical value ( $d_U$ ): **1.56**

The decision rule states that if  $d < d_L$ , we reject the null hypothesis and conclude positive autocorrelation exists. If  $d > d_U$ , we fail to reject the null hypothesis. If  $d$  falls between  $d_L$  and  $d_U$ , the result is inconclusive. Since our test statistic of **1.3475** falls within the inconclusive range ( $0.86 \leq 1.3475 \leq 1.56$ ), we technically cannot draw a definitive conclusion regarding positive autocorrelation. However, since the value does not fall into the strong rejection zone ( $d < 0.86$ ), we state that we do not have sufficient evidence to definitively reject the null hypothesis.

In simple terms, based on the non-rejection of  $H_0$ , we conclude there is no statistically significant first-order autocorrelation among the residuals, suggesting the model's assumption of independent errors is reasonably met.

## What to Do if Significant Autocorrelation is Detected

If you reject the null hypothesis and conclude that significant autocorrelation is present in the residuals, the validity of your standard error estimates is compromised. Addressing this is crucial for accurate statistical inference. Here are several strategic options to correct or mitigate the problem:

For positive serial correlation ( $d$  near 0), consider restructuring the model by adding lagged variables. This often means including prior values of the dependent variable (e.g.,  $Y_{t-1}$ ) and/or prior values of the independent variable to the model to capture the time-dependent structure the original model missed.

For negative serial correlation ( $d$  near 4), which is less common, check to make sure that none of your variables are overdifferenced. Excessive differencing of time series data can artificially induce negative correlation in the residuals.

For seasonal correlation, which the Durbin-Watson test may not catch (as it tests for first-order only), consider implementing diagnostic tests for higher orders (like the Breusch-Godfrey test). If seasonal patterns are confirmed, consider adding seasonal dummy variables (e.g., for months or quarters) to the model to account for recurring periodic effects.

## Conclusion: Ensuring Model Validity

The Durbin-Watson test is an essential diagnostic tool that ensures the integrity of time-series regression analysis by verifying the independence of errors. By systematically performing the calculation using the Data Analysis ToolPak and interpreting the resulting statistic against critical values, analysts can confidently determine if model assumptions are met or if structural refinements are necessary.