

How do I calculate the Residual Sum of Squares in Excel?

Authored by
stats writer

April 24, 2024

RECOMMENDED CITATION

stats writer (2024). *How do I calculate the Residual Sum of Squares in Excel?*.

PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=138696>

Calculating the Residual Sum of Squares (RSS) in Excel is a method used to measure the accuracy of a regression model. It is the sum of the squared differences between the predicted values and the actual values. To calculate the RSS in Excel, first, input the predicted values and the actual values in two separate columns. Then, use the formula "`=SUMXMY2(predicted values, actual values)`" to calculate the squared differences. Finally, sum up all the squared differences to get the RSS value. This method is commonly used in statistical analysis to evaluate the effectiveness of a regression model and make necessary adjustments.

Calculate Residual Sum of Squares in Excel

A is the difference between an observed value and a predicted value in a regression model.

It is calculated as:

Residual = Observed value - Predicted value

One way to understand how well a regression model fits a dataset is to calculate the residual sum of squares, which is calculated as:

Residual sum of squares = $\sum(e_i)^2$

where:

Σ : A Greek symbol that means "sum" e_i : The i th residual

The lower the value, the better a model fits a dataset.

This tutorial provides examples of how to calculate the residual sum of squares for a simple linear regression model and a multiple linear regression model in Excel.

Example 1: Residual Sum of Squares for Simple Linear Regression

Suppose we have the following dataset in Excel:

	A	B	C	D	E	F
1	x	y				
2	6	16				
3	7	17				
4	7	17				
5	8	19				
6	12	22				
7	14	24				
8	15	25				
9	16	27				
10	16	33				
11	19	34				
12	21	35				
13	22	35				
14	23	34				
15	24	36				
16	24	39				
17						
18						
19						
20						
21						
22						
23						
24						
25						

To calculate the residual sum of squares for a simple linear regression model using x as the predictor variable and y as the response variable we can use the

function, which uses the following syntax:

LINEST(known_ys, , ,)

where:

known_ys: The range of y-values
known_sx: The range of x-values
const: Whether to force the constant b to be zero. We will leave this blank.
stats: A list of regression statistics. We will specify this to be TRUE.

The following screenshot shows how to use this function in practice:

	A	B	C	D	E	F	G	H	I
1	x	y							
2	6	16		1.202164	8.779581		=LINEST(B2:B16,A2:A16,,TRUE)		
3	7	17		0.081237	1.366128				
4	7	17		0.943963	1.975904				
5	8	19		218.9896	13				
6	12	22		854.9788	50.75456				
7	14	24							
8	15	25							
9	16	27							
10	16	33							
11	19	34							
12	21	35							
13	22	35							
14	23	34							
15	24	36							
16	24	39							
17									
18									
19									
20									
21									
22									
23									

The residual sum of squares for the regression model is displayed in the last cell of the second column of the output. In this example, the residual sum of squares turns out to be 50.75.

Example 2: Residual Sum of Squares for Multiple Linear Regression

Suppose we have the following dataset in Excel:

	A	B	C	D	E	F	G
1	x1	x2	y				
2	6	8	16				
3	7	9	17				
4	7	11	17				
5	8	11	19				
6	12	11	22				
7	14	12	24				
8	15	12	25				
9	16	12	27				
10	16	13	33				
11	19	13	34				
12	21	14	35				
13	22	15	35				
14	23	15	34				
15	24	17	36				
16	24	19	39				
17							
18							
19							
20							
21							
22							
23							
24							
25							
26							
27							

Once again we can use the **LINEST()** function to calculate the residual sum of squares for the model.

The only difference is that we'll specify two columns of values for the **known_xs** argument:

	A	B	C	D	E	F	G	H	I	J	K
1	x1	x2	y								
2	6	8	16		0.22402	1.11098	7.334584		=LINEST(C2:C16, A2:B16, , TRUE)		
3	7	9	17		0.475237	0.210801	3.373707				
4	7	11	17		0.944982	2.037806	#N/A				
5	8	11	19		103.0548	12	#N/A				
6	12	11	22		855.9015	49.83183	#N/A				
7	14	12	24								
8	15	12	25								
9	16	12	27								
10	16	13	33								
11	19	13	34								
12	21	14	35								
13	22	15	35								
14	23	15	34								
15	24	17	36								
16	24	19	39								
17											
18											
19											
20											
21											
22											
23											
24											

The residual sum of squares for this multiple linear regression model turns out to be 49.83.