How do I calculate standardized residuals in Excel?

Authored by

stats writer

April 23, 2024

ARABPSYCHOLOGY.COM

RECOMMENDED CITATION

stats writer (2024). *How do I calculate standardized residuals in Excel?*. PSYCHOLOGICAL SCALES. Retrieved from https://scales.arabpsychology.com/?p=138533

The process of calculating standardized residuals in Excel involves taking the difference between the observed data values and the predicted values, and then dividing it by the standard deviation of the residuals. This results in a standardized measure of the difference between the actual data and the predicted values. To do this in Excel, you will need to use the appropriate formulas and functions, such as the "STDEV.S" function and the "RESID" function. By following these steps, you can easily calculate and interpret standardized residuals to assess the accuracy of your predicted values.

Calculate Standardized Residuals in Excel

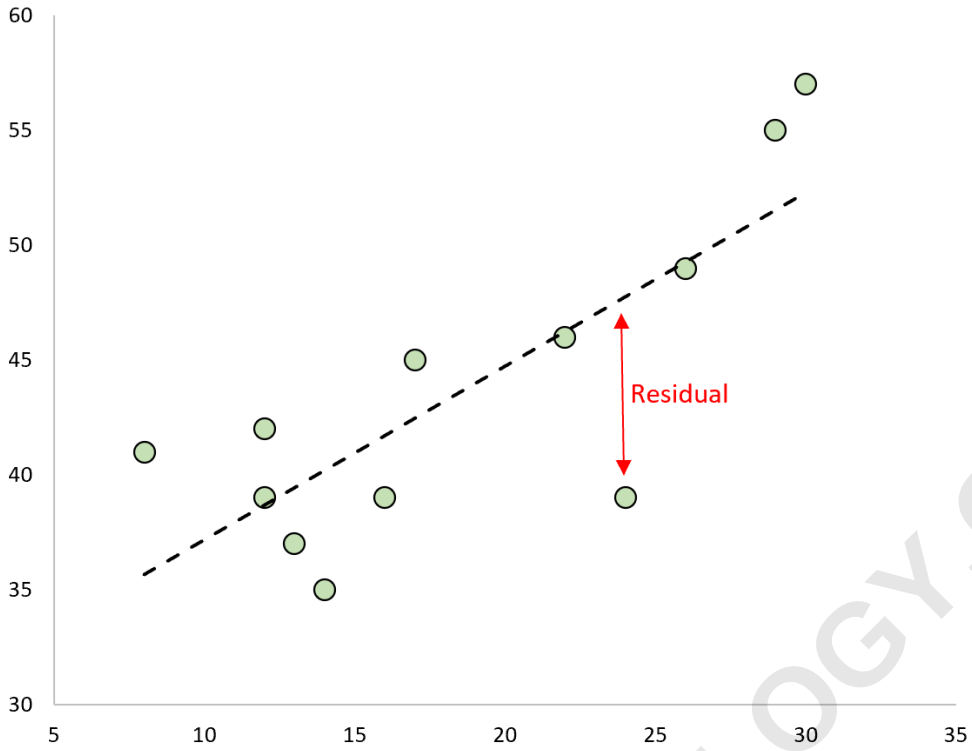
A residual is the difference between an observed value and a predicted value in a **regression model**.

It is calculated as:

Residual = Observed value - Predicted value

If we plot the observed values and overlay the fitted regression line, the residuals for each would be the vertical distance between the observation and the regression line:

ARABPSYCHOLOGY.COM



ARABPSYCHOLOGY.COM

One type of residual we often use to identify outliers in a regression model is known as a standardized residual.

It is calculated as:

ri = ei / s(ei) = ei / RSE√1-hii

where:

ei: The ith residualRSE: The residual standard error of the modelhii: The leverage of the ith observation

In practice, we often consider any standardized residual

with an absolute value greater than 3 to be an outlier.

This tutorial provides a step-by-step example of how to calculate standardized residuals in Excel.

Step 1: Enter the Data

First, we'll enter the values for a small dataset into Excel:

	A	B	C	D	E	F	G
1		X	Y				
2		8	41				
3		12	42				
4		12	39				
5		13	37				
6		14	35				
7		16	39				
8		17	45				
9		22	46				
10		24	39				
11		26	49				
12		29	55				
13		30	57				
14							
15							
16							
17							
18							
19							
20							
21							
22							
23							
24							

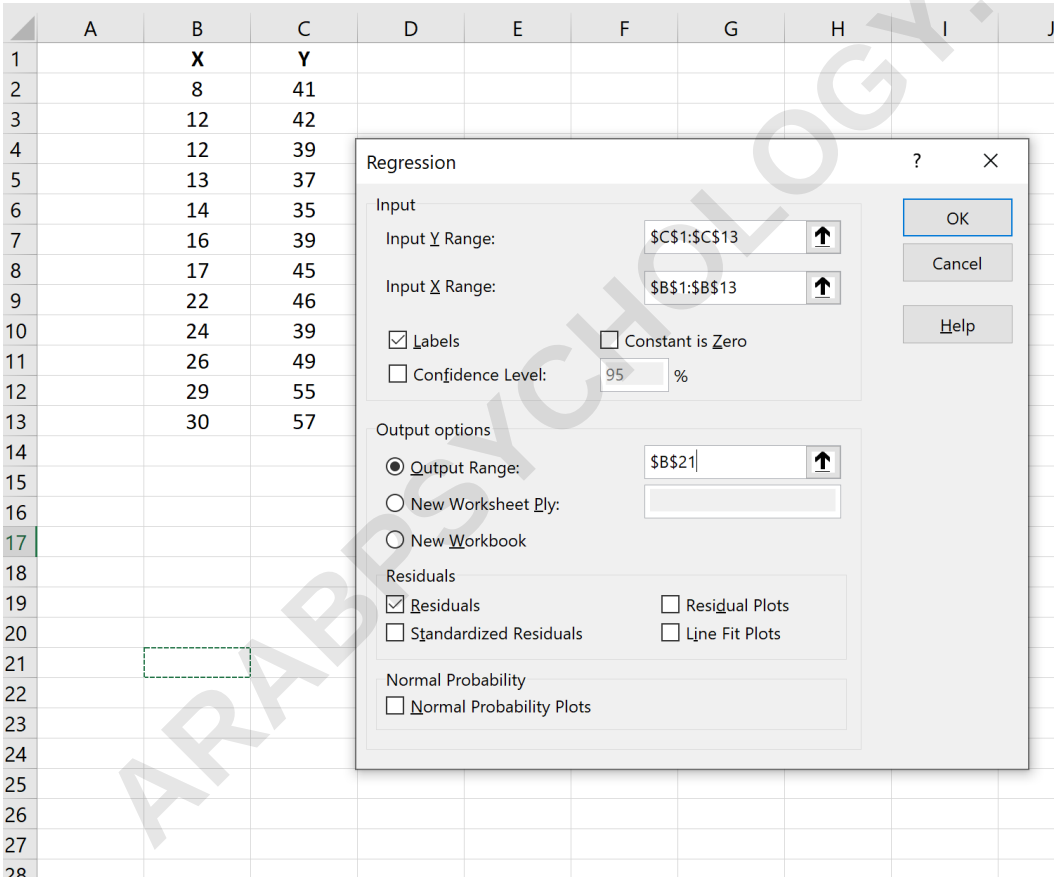
ARABPSYCHOLOGY.COM

Step 2: Calculate the Residuals

Next, we'll go to the Data tab along the top ribbon and click Data Analysis within the Analysis group:

If you haven't installed this Add-in already, check out on how to do so. It's easy to install and completely free.

Once you've clicked Data Analysis, click the option that says Regression and then click OK. In the new window that pops up, fill in the following information and click OK:



The residual for each observation will appear in the output:

ARABPSYCHOLOGY.COM

RESIDUAL OUTPUT

Observation	Predicted Y	Residuals
1	35.6732	5.326796
2	38.69435	3.305645
3	38.69435	0.305645
4	39.44964	-2.44964
5	40.20493	-5.20493
6	41.71551	-2.71551
7	42.47079	2.529206
8	46.24723	-0.24723
9	47.75781	-8.75781
10	49.26838	-0.26838
11	51.53425	3.465752
12	52.28954	4.710464

ARABPSYCHOLOGY.COM

Copy and paste these residuals in a new column next to the original data:

	A	B	C	D	E
1		**X**	**Y**	**Residuals**	
2		8	41	5.326796	
3		12	42	3.305645	
4		12	39	0.305645	
5		13	37	-2.44964	
6		14	35	-5.20493	
7		16	39	-2.71551	
8		17	45	2.529206	
9		22	46	-0.24723	
10		24	39	-8.75781	
11		26	49	-0.26838	
12		29	55	3.465752	
13		30	57	4.710464	
14					
15					
16					
17					
18					
19					
20					

Step 3: Calculate the Leverage

Next, we need to calculate the leverage of each observation.

The following image shows how to do so:

ARABPSYCHOLOGY.COM

	A	B	C	D	E	F
1		**X**	**Y**	**Residuals**	**Leverage**	
2		8	41	5.326796	0.271607	
3		12	42	3.305645	0.156184	
4		12	39	0.305645	0.156184	
5		13	37	-2.44964	0.135733	
6		14	35	-5.20493	0.118644	
7		16	39	-2.71551	0.094551	
8		17	45	2.529206	0.087547	
9		22	46	-0.24723	0.102956	
10		24	39	-8.75781	0.132652	
11		26	49	-0.26838	0.175795	
12		29	55	3.465752	0.265723	
13		30	57	4.710464	0.302423	
14	**n**	12				
15	**Mean**	18.583				
16	**SS**	594.917				
17						
18						
19						
20						

ARABPSYCHOLOGY.COM

Here are the formulas used in the various cells:

B14: =COUNT(B2:B13)B15: =AVERAGE(B2:B13)B16: =DEVSQ(B2:B13) E2: =1/$B$14+(B2-$B$15)^2/$B$16

Step 4: Calculate the Standardized Residuals

Lastly, we can calculate the standardized residuals using the formula:

ri = ei / RSE√1-hii

The RSE for the model can be found in the model output from earlier. It turns out to be 4.44:

13		30	57	4.710464	0.302423	
14	**n**	12				
15	**Mean**	18.583				
16	**SS**	594.917				
17						
18						
19						
20						
21		SUMMARY OUTPUT				
22						
23		*Regression Statistics*				
24		Multiple R	0.795222			
25		R Square	0.632378			
26		Adjusted R Square	0.595615			
27		Standard Error	4.441741			
28		Observations	12			
29						
30		ANOVA				
31			*df*	*SS*	*MS*	*F*
32		Regression	1	339.376	339.376	17.2018
33		Residual	10	197.2907	19.72907	
34		Total	11	536.6667		
35						
36			*Coefficients*	*tandard Erro*	*t Stat*	*P-value*
37		Intercept	29.6309	3.618911	8.187794	9.6038
38		X	0.755288	0.182106	4.147509	0.00198
39						

Thus, we can use the following formula to calculate the standardized residual for each observation:

ARABPSYCHOLOGY.COM

	A	B	C	D	E	F
1		**X**	**Y**	**Residuals**	**Leverage**	**Standardized Residuals**
2		8	41	5.326796	0.271607	=D2/(4.44*SQRT(1-E2))
3		12	42	3.305645	0.156184	0.810
4		12	39	0.305645	0.156184	0.075
5		13	37	-2.44964	0.135733	-0.593
6		14	35	-5.20493	0.118644	-1.249
7		16	39	-2.71551	0.094551	-0.643
8		17	45	2.529206	0.087547	0.596
9		22	46	-0.24723	0.102956	-0.059
10		24	39	-8.75781	0.132652	-2.118
11		26	49	-0.26838	0.175795	-0.067
12		29	55	3.465752	0.265723	0.911
13		30	57	4.710464	0.302423	1.270
14	**n**	12				
15	**Mean**	18.583				
16	**SS**	594.917				
17						
18						
19						
20						

ARABPSYCHOLOGY.COM

From the results we can see that none of the standardized residuals exceed an absolute value of 3. Thus, none of the observations appear to be outliers.

It's worth noting in some cases that researchers consider observations with standardized residuals that exceed an absolute value of 2 to be considered outliers.

It's up to you to decide whether to use an absolute value of 2 or 3 as the threshold for outliers, depending on the specific problem you're working on.

What Are Standardized Residuals?