

Finding Residuals

Learning Outcomes

- Given a Regression line and a data point, find the residual

In the linear regression part of statistics we are often asked to find the residuals. Given a data point and the regression line, the residual is defined by the vertical difference between the observed value of y and the computed value of \hat{y} based on the equation of the regression line:

$$\text{Residual} = y - \hat{y}$$

Example 1

A study was conducted asking female college students how tall they are and how tall their mother is. The results are shown in the table below:

Table of Mother and Daughter Heights

Mother's Height	63	67	64	60	65	67	59	60
Daughter's Height	58	64	65	61	65	67	61	64

The equation of the regression line is

$$\hat{y} = 30.28 + 0.52x$$

Find the residual for the mother who is 59 inches tall.

Solution

First note that the Daughter's Height associated with the mother who is 59 inches tall is 61 inches. This is y . Next we use the equation of the regression line to find \hat{y} . Since $x = 59$, we have

$$\hat{y} = 30.28 + 0.52(59)$$

We can use a calculator to get:

$$\hat{y} = 60.96$$

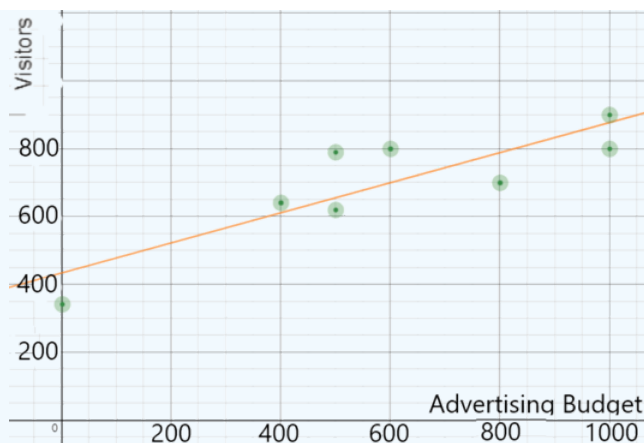
Now we are ready to put the values into the residual formula:

$$\text{Residual} = y - \hat{y} = 61 - 60.96 = 0.04$$

Therefore the residual for the 59 inch tall mother is 0.04. Since this residual is very close to 0, this means that the regression line was an accurate predictor of the daughter's height.

Example 2

An online retailer wanted to see how much bang for the buck was obtained from online advertising. The retailer experimented with different weekly advertising budgets and logged the number of visitors who came to the retailer's online site. The regression line for this is shown below.



Find the residual for the week when the retailer spent \$600 on advertising.

Solution

First notice that the point of the scatterplot with x-coordinate of 600 has y-coordinate 800. Thus $y = 800$. Next note that the point on the line with x-coordinate 600 has y-coordinate 700. Thus $\hat{y} = 700$. Now we are ready to put the values into the residual formula:

$$\text{Residual} = y - \hat{y} = 800 - 700 = 100$$

Therefore the residual for the \$600 advertising budget is 100.

Exercise

Data was taken from the recent Olympics on the GDP in trillions of dollars of 8 of the countries that competed and the number of gold medals that they won. The equation of the regression line is:

$$\hat{y} = 7.55 + 1.57x$$

The table below shows the data:

GDP	21	1.6	16	1.8	4	5.4	3.1	2.3
Medals	46	8	26	19	17	12	10	9

Find the residual for the country with a GDP of 4 trillion dollars.

- [Calculating residual example | Exploring bivariate numerical data | AP Statistics | Khan Academy](#)
- [Finding a Residual](#)

This page titled [Finding Residuals](#) is shared under a [CC BY 4.0](#) license and was authored, remixed, and/or curated by [Larry Green](#).