



Lesson 17: Analyzing Residuals

Student Outcomes

- Students use a graphing calculator to construct the residual plot for a given data set.
- Students use a residual plot as an indication of whether the model used to describe the relationship between two numerical variables is an appropriate choice.

Lesson Notes

In this lesson, students use technology to generate residual plots. Students build on their knowledge of using technology to determine the least squares line for a data set. The steps for using TI-83/84 graphing calculators to create residual plots are provided in the teacher notes, which can be printed and distributed for students to follow during the lesson as a guide. Similar summaries may be needed for any graphing software or other graphing calculators that students may use. Residual plots are also analyzed to determine the appropriateness of informally fitting the linear models to the data set.

Classwork

Example 1 (10 minutes): Predicting the Pattern in the Residual Plot

MP.2

Students learn how to use the pattern in the scatter plot to predict what the residual plot will look like. Vertical distances from the least squares line in the scatter plot are plotted as the y -values in the residual plot, with the x -coordinates remaining the same. Note that the scales on the vertical axes of the two graphs are different, and as a result, the sizes of the displacements are greatly exaggerated in the residual plot.

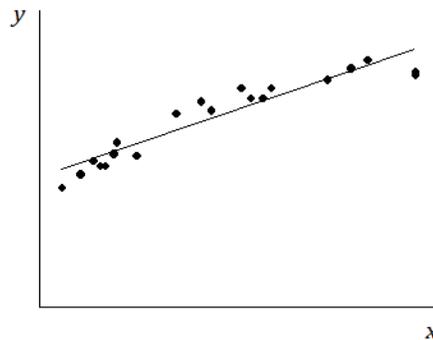
Ask students:

- What will the residual plot look like?

If students have trouble with this, explain each point in the scatter plot one by one. For each point, look at whether it has a positive or negative residual and whether the residual is large or small relative to the other residuals. Once they have seen two or three points in the scatter plot and how they translate to points in the residual plot, students should see how the pattern translates as a whole. Students look at the residual plot as an indication of the fit of the points to the line.

Example 1: Predicting the Pattern in the Residual Plot

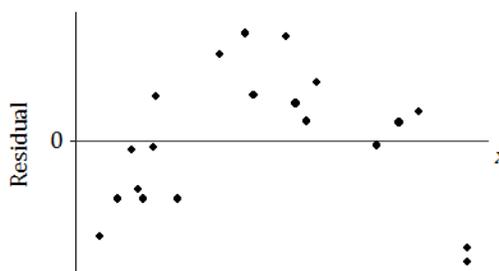
Suppose you are given a scatter plot and least squares line that looks like this:



Describe what you think the residual plot would look like.

If students struggle, have them look carefully at the scatter plot. Point out that moving from left to right, the points initially tend to be below the least squares line, then move above it, and then below it. The residuals are negative, then positive, and then negative again. This means that the points in the residual plot will be below the horizontal axis, then above it, and then below it again.

The residual plot has an arch shape like this:



Why is looking at the pattern in the residual plot important?

Discuss with students that the residual plot provides another indication of the type of relationship between two variables.

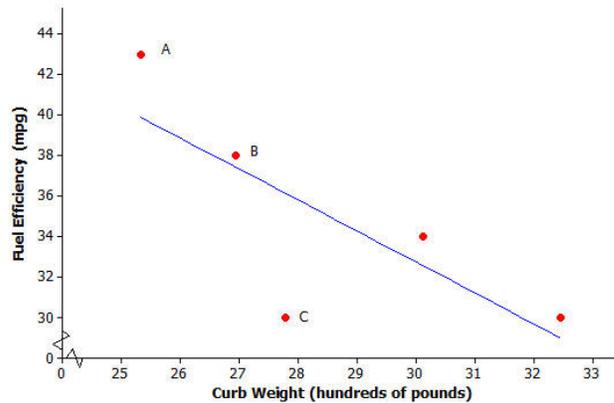
Example 2 (10 minutes): The Meaning of Residuals

The purpose of this example is to remind students of what a residual plot is before they construct one using a calculator. If students seem comfortable with the concept of a residual plot following the previous lesson and Problem Set, this example can be covered quickly. If students need some review, this is a good opportunity to do that.

Example 2: The Meaning of Residuals

Suppose that you have a scatter plot and that you have drawn the least squares line on your plot. Remember that the residual for a point in the scatter plot is the vertical distance of that point from the least squares line.

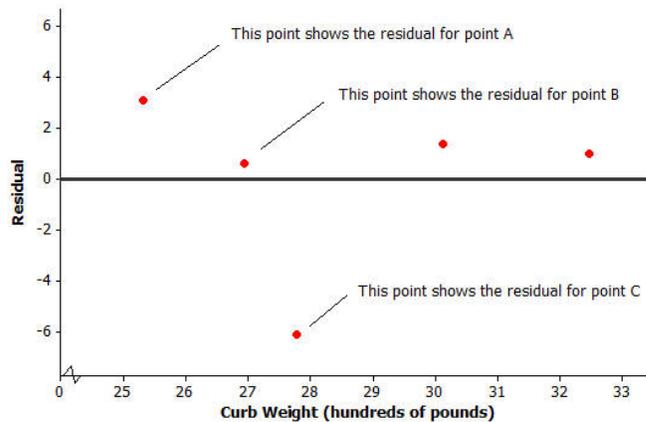
In the previous lesson, you looked at a scatter plot showing how fuel efficiency was related to curb weight for five compact cars. The scatter plot and least squares line are shown below.



Consider the following questions:

- What kind of residual does Point A have?
Point A has a large positive residual.
- What kind of residual does Point B have?
Point B has a small positive residual.
- What kind of residual does Point C have?
Point C has a very large negative residual.

You also looked at the residual plot for this data set:



Your teacher will now show how to use a graphing calculator or graphing program to construct a scatter plot and a residual plot. Consider the following example.

Example 3 (15 minutes): Using a Graphing Calculator to Construct a Residual Plot

MP.5

Even though students have already learned how to calculate the equation of the least squares line using a graphing calculator, the instructions are repeated here for completeness. In this example, students construct a scatter plot with the least squares line and construct a residual plot using technology.

The instructions shown here are for the TI-83 and TI-84 calculators and are intended as a guide only. Alternate graphing software may also be used. Most graphing programs have similar directions to those listed for the TI-83 or TI-84.

Note: Students using the TI-83 or TI-84 should clear out any equations under $Y =$ before beginning this example.

Example 3: Using a Graphing Calculator to Construct a Residual Plot

In an earlier lesson, you looked at a data set giving the shoe lengths and heights of 12 adult women. This data set is shown in the table below.

x (Shoe Length)	y (Height)
inches	inches
8.9	61
9.6	61
9.8	66
10.0	64
10.2	64
10.4	65
10.6	65
10.6	67
10.5	66
10.8	67
11.0	67
11.8	70

Use a calculator to construct the scatter plot (with least squares line) and the residual plot for this data set.

Calculation of the equation of the least squares line:

1. From the home screen, go to the statistics editor by pressing STAT, ENTER.
2. Enter the x -values into L_1 and the y -values into L_2 .
3. Press 2nd, QUIT to return to the home screen.
4. Press STAT, select CALC, select $\text{LinReg}(a + bx)$, and press ENTER.
5. Type L_1, L_2, Y_1 so that your entry reads $\text{LinReg}(a + bx) L_1, L_2, Y_1$. (Y_1 is accessed by pressing VARS and selecting Y-VARS, Function, 1.) Press ENTER.
6. The y -intercept, a , and the slope, b , of the least squares line are displayed on the screen. (Also, the calculator stores in its memory a list called RESID, which is used in constructing the residual plot. Furthermore, since Y_1 was included in the entry, the equation of the least squares line is entered as the Y_1 equation in the Y = menu.)

Construction of scatter plot:

1. From the home screen press 2nd, STAT PLOT.
2. Select Plot1, and press ENTER.
3. Select *On*. Under *Type* choose the first (scatter plot) icon, for Xlist enter L_1 , for Ylist enter L_2 , and under *Mark* choose the first (square) symbol.
4. Press 2nd, QUIT to return to the home screen.
5. Press Y =.
6. Go to any unwanted graph equations and press CLEAR. Make sure that only Plot1 is selected (not Plot2 or Plot3).
7. Press Zoom, select ZoomStat (option 9), and press ENTER.
8. The scatter plot and the least squares line are displayed.

Construction of residual plot:

1. From the home screen, press 2nd, STATPLOT.
2. Select Plot2, and press ENTER.
3. Select *On*. Under *Type*, choose the first (scatter plot) icon, for Xlist enter L_1 , for Ylist enter RESID, and under *Mark* choose the first (square) symbol. (RESID is accessed by pressing 2nd, LIST, selecting NAMES, scrolling down to RESID and pressing ENTER.)
4. Press 2nd, QUIT to return to the home screen.
5. Press Y =.
6. First, deselect the equation of the least squares line in Y_1 by going to the = sign for Y_1 and pressing ENTER. Then deselect Plot1, and make sure that Plot2 is selected.
7. Press Zoom, select ZoomStat (option 9), and press ENTER.
8. The residual plot is displayed.

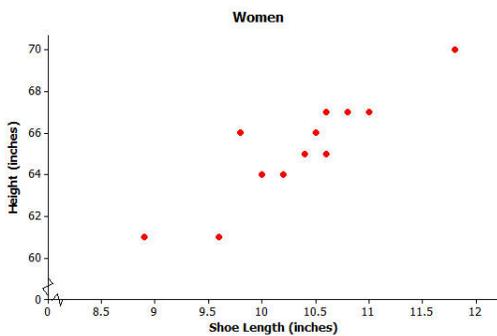
Note that when constructing a residual plot, you need the correct set of residuals to form the list called *RESID*. This is only the case if you have run the $\text{LinReg}(a + bx)$ or $\text{LinReg}(ax + b)$ command for the data set on which you are working.

Students often ask why we use $\text{LinReg}(a + bx)$ rather than $\text{LinReg}(ax + b)$. Explain how the two forms offer different representations of a linear equation. The more common convention in statistics is that a is the y -intercept and b is the slope, not vice-versa.

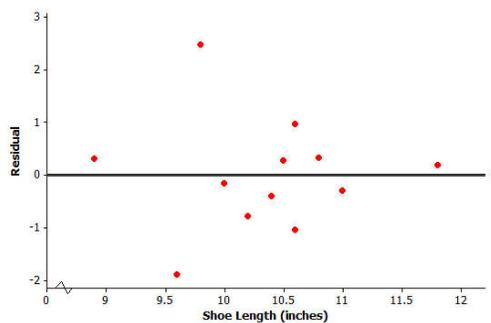
To save time and help students do the problem set, it might be worthwhile to give each student a copy of the calculator instructions.

The equation of the least squares line for this example is $y = 32.679 + 3.1469x$, where x is the shoe length (in inches), and y is the height (in inches). The scatter plot and residual plot are shown below.

Scatter plot:



Residual plot:



Closing (5 minutes)

Discuss the Lesson Summary with students.

Lesson Summary

- After fitting a line, the residual plot can be constructed using a graphing calculator.
- A pattern in the residual plot indicates that the relationship in the original data set is not linear.

Exit Ticket (5 minutes)

Exit Ticket Sample Solutions

1. If you see a random scatter of points in the residual plot, what does this say about the original data set?

A random scatter of points in the residual plot tells us that a linear function is an appropriate model for the relationship between the two variables in the original data set.

2. Suppose a scatter plot of bivariate numerical data shows a linear pattern. Describe what you think the residual plot would look like. Explain why you think this.

Because the relationship between the two variables is linear, the residual plot would have a random pattern.

Problem Set Sample Solutions

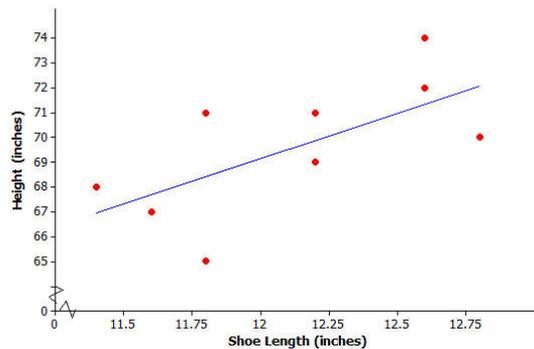
Assist those students still needing help with the calculator or statistical program. If students are using a TI graphing calculator (e.g., a TI-83 or TI-84), tell them to turn off the STAT plot for 2. The first question asks only about the least squares line and the scatter plot.

Consider again a data set giving the shoe lengths and heights of 10 adult men. This data set is shown in the table below.

x (Shoe Length)	y (Height)
inches	inches
12.6	74
11.8	65
12.2	71
11.6	67
12.2	69
11.4	68
12.8	70
12.2	69
12.6	72
11.8	71

- Use your calculator or graphing program to construct the scatter plot of this data set. Include the least squares line on your graph. Explain what the slope of the least squares line indicates about shoe length and height.

The slope indicates that as the shoe length increases, so does the height. The following illustrates the scatter plot and residual plot from a calculator. (Note: If it is not possible for you to observe your students' work on their calculators, have them show you a sketch of each graph.)



- Use your calculator to construct the residual plot for this data set.

See plot below.

- Make a sketch of the residual plot on the axes given below. Does the scatter of points in the residual plot indicate a linear relationship in the original data set? Explain your answer.

There is no curved pattern in the residual plot, and the points in the residual plot appear to be scattered at random around the zero line. This indicates that the relationship in the original data set is linear.

