

## 14.2: Linear Fit Demo

### Learning Objectives

- State the relationship between  $MSE$  and fit

### Instructions

This demonstration allows you to explore fitting data with linear functions. When the demonstration begins, five points are plotted in the graph. The  $X$  axis ranges from 1 to 5 and the  $Y$  axis ranges from 0 to 5. The five points are plotted in different colors; next to each point is the  $Y$  value of that point. For example, the red point has the value 1.00 next to it. A vertical black line is drawn with the  $Y$  value of 3.0; this line consists of the predicted values for  $Y$ . (It is clear that this line does not contain the best predictions.) This line is called the "regression line." The equation for the regression line is  $Y' = 0X + 3$  where  $Y'$  is the predicted value for  $Y$ . Since the slope is 0, the same prediction of 3 is made for all values of  $X$ .

The error of prediction for each point is represented by a vertical line between the point and the regression line. For the point with a value of 1 on the  $X$  axis, the line goes from the point (1, 1) to the point (1, 3) on the line. The length of the vertical line is 2. This means the error of prediction is 2 and the squared error of prediction is  $2 \times 2 = 4$ . This error is depicted graphically by the graph on the right. The height of the red square is the error of prediction and the area of the red square is the squared error of prediction.

The errors associated with the other points are plotted in a similar way. Therefore the height of the stacked squares is the sum of the errors of prediction (the lengths of the lines are used, so all errors are positive) and the area of all squares is the total sum of squared errors.

This demonstration allows you to change the regression line and examine the effects on the errors of prediction. If you click and drag an end of the line, the slope of the line will change. If you click and drag the middle of the line, the intercept will change. As the line changes, the errors and squared errors are updated automatically.

You can also change the values of the points by clicking on them and dragging. You can only change the  $Y$  values.

You can get a good feeling for the regression line and error by changing the points and the slope and intercept of the line and observing the results.

To see the line that minimizes the squared errors of prediction click the "OK" button.

1. Notice that the total deviation is 6. Compute this by summing the 5 separate absolute deviations ( $|1 - 3| + |2 - 3|$  etc.)
2. The total area is 10. Compute this by summing the squared deviations.
3. Click middle of the black line and drag it so that it is at 4. Has the error increased or decreased?
4. Click the left end of the line and drag it until the intercept is about 1. How has this affected the error?
5. Drag the line further so it has an intercept of about 0. Then drag the upper portion of the line so that it goes through the point at 5.0. Within rounding error, the error should be 0 and the equation for the line should be  $Y' = 1X + 0$ .
6. Click on the green point at 3.0 and drag it down so that its value is about 2.0. Notice the error now is all based on this one point.
7. Adjust the line to see if you can make the absolute error any smaller.
8. Adjust the line to see if you can make the squared error (area) as small as you can and note equation of the line and the size of the area. Press the OK button to see the smallest possible area. How does the line compare with the one you chose?
9. Move all the points around and then try again to find the line that makes the area smallest. Compare the area to the line that gives the smallest area.
10. After you have found the line that gives the smallest area (smallest squared error) see if you can change the line so that the absolute error is lower than for this line that gives the smallest area.

### Illustrated Instructions

#### Video Demo

The demonstration starts by dragging each of the 5 points to different locations on the  $Y$  axis. Notice how these changes influence the total deviation and area. The video continues by repositioning the regression line by dragging either end as well as the middle. Finally the regression line that minimizes the squared errors is found by clicking the "OK" button.

This page titled [14.2: Linear Fit Demo](#) is shared under a [Public Domain](#) license and was authored, remixed, and/or curated by [David Lane](#) via [source content](#) that was edited to the style and standards of the LibreTexts platform.