Squared Differences Simulation*

David Lane

This work is produced by OpenStax-CNX and licensed under the
Creative Commons Attribution License 1.0†

Begin by answering the questions, even if you have to guess. The first time you answer the questions you will not be told whether you are correct or not.

Once you have answered all the questions, answer them again using the simulation to help you. This time you will get feedback about each individual answer.

Show Simulation (Section 1: General Instructions)

This is a Java Applet. To view, please see http://cnx.org/content/m11209/latest/

1 General Instructions

This demonstration allows you to examine the sum of absolute deviations from a given value. The graph to the left shows the numbers 1, 2, 3, 4, and 5 and their deviations from 0. The first number, 1, is represented by a red dot. The deviation from 0 is represented by a red line from the red dot to the black line. The value of the black line is 0.

Similarly, the number 2 is represented by a blue dot and its deviation from 0 is represented by a blue line. The graph with the colored rectangles shows the sum of the absolute deviations. Since the deviations for the numbers 1, 2, 3, 4, and 5 from 0 are the numbers themselves, the sum of the deviations is equal to 1 + 2 + 3 + 4 + 5 = 15 as shown by the height of the colored bar.

In this demonstration, you can move the black bar by clicking on it and dragging it up or down. To see how it works, move it up to 1.0. The deviation of the red point from the black bar is now 0 since they are both 1. The sum of the deviations is now 10.

As you move the bar up and down, the value of the sum of absolute deviations changes. See if you can find the placement of the black bar that produces the smallest value for the sum of the absolute deviations. To check and see if you found the smallest value, click the "OK" button at the bottom of the graph. It will move the bar to the location that produces the smallest sum of absolute deviations.

You can also move the individual points. Click on one of the points and move it up or down and note the effect.

Your goal for this demonstration is to discover a rule for determining what value will give you the smallest sum of absolute deviations.

When you have discovered the rule, go back and answer the questions again.

2 Step by Step Instructions

Show Questions (p. 1)

If it is not already there, move the black bar at the bottom of the graph up so that it crosses the Y axis at 1. The bar should go right through the red circle. Notice the numerical indicator of the black bar immediately to its right.

*Version 1.3: Jul 14, 2003 5:43 pm -0500

†http://creativecommons.org/licenses/by/1.0

The deviation of the red circle from the bar is 0, so you won't see a red rectangle on the right-hand portion of the graph. The line between the bar and the blue circle is the deviation of the circle from the bar. It has a length of 1. Notice that the height of the blue rectangle is 1.

The green line has a length of 2 and the height of the green rectangle is also 2. The total height of the rectangle is the sum of all the line lengths: $0 + 1 + 2 + 3 + 4 = 10$. This height is the sum of the absolute deviations from the bar. It is marked below the rectangle.

Your goal is to find the placement of the bar that gives you the shortest rectangle. This will be the value that minimizes the sum of the rectangles. Move the bar up and down until you think you have found this value. Then, to make sure you are correct, click on the "OK" button at the bottom of the graph. This will move the black bar to the correct location. If nothing changes, you found the correct location on your own.

Now, change the value of the green circle from 3 to somewhere between 2 and 3. You move the circle by clicking on it and dragging it. Notice that the value of the point is shown at the bottom of the graph in green.

Next, find the value that minimizes the sum of absolute deviations for the new data. Once again it is the same value as the green circle.

Now move the blue circle to somewhere between 3 and 4 and again find the value that minimizes the sum of absolute differences. This time it is the value of the blue circle.

How do you know which point it will be? The correct placement of the bar will always be at the value of the circle with the middle value. That is, the circle that has two point higher than it and two points lower than it.

Why is this? If the bar is at the circle with the middle value, then moving the bar will bring the bar closer to two points but farther from three points. So, any movement of the bar from the middle value increases the sum of absolute deviations.

This is a Java Applet. To view, please see http://cnx.org/content/m11209/latest/

3 Summary

The middle number minimizes the sum of the absolute deviations. If you have 7 numbers, say 0, 1, 2, 5, 6, 9, and 12, then the middle number is the fourth highest which is 5. Therefore 5 is the value that minimizes the sum of the absolute deviations. This middle number is called the median.

If you have an even number of numbers such as the four numbers: 2, 4, 6, and 9, then any value between 4 and 6 will give you the same minimum sum of the absolute deviations. By convention, we define the median as the average of the two numbers closest the middle (2 and 4 in this case). So the median of the numbers 2, 4, 6, and 9 is 5.