Modeling Data: Lines of Best Fit

Our world is filled with data as persons in such areas as government, health, social sciences, sciences, education, and business gather information that can be used to predict population growth, determine causes of health problems, or define marketing goals. The advent of the computer has made a wealth of data available, as well as providing ways to analyze the data easily. In an earlier section, we plotted data that gives the balance of trade that the United States had with China for the years 1987 – 1994 and we made a scatterplot of this data. (see table and plot below)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Balance of Trade ($ Billion)</td>
<td>-2.8</td>
<td>-3.5</td>
<td>-6.2</td>
<td>-10.4</td>
<td>-12.7</td>
<td>-18.3</td>
<td>-22.8</td>
<td>-29.5</td>
</tr>
</tbody>
</table>

Notice that the points suggest a line. By writing an equation for the line that comes close to all these points we can approximate the relation between these two sets of data with a mathematical model. In this case we will use a linear equation to model the data called the “line of best fit.”

Example 1:

a) Although your calculator will calculate the values for the slope and $y$-intercept that model the line most accurately, use two points on the line to write an equation for a line that roughly approximates the data.

b) Use the equation to predict the balance of trade with China in the year 1998.

c) According to your model, when will the balance of trade between the United States and China reach a deficit of $50\ billion$?

d) According to the actual points on the plot, does it appear that the deficit will reach $50\ billion$ before or after the year you predicted?

e) From the line of best fit that you calculated, what is the yearly rate of change of the deficit?
Solution:

a) The line appears to go through the points $(1, -3.5)$ and $(7, -22.8)$ so we will use these point to write our equation. 
\[ m = \frac{-3.5 - (-22.8)}{1 - 7} = \frac{19.3}{-6} = -3.2167. \]
Using point-slope, we approximate our model to be 
\[ y - (-3.5) = -3.2(x - 1) \Rightarrow y + 3.5 = -3.2x + 3.2 \Rightarrow y = -3.2x - 0.3, \]
where $y$ is the deficit for $x$ years since 1987.

b) 1998 is 11 years since 1987 \( \Rightarrow x = 11 \Rightarrow y = -3.2(11) - 0.3 = -35.5 \). In 1998 the trade deficit with China will reach $35.5 \text{ billion}$ according to our model.

c) Deficit of $50 \text{ billion} \Rightarrow y = -50 \Rightarrow -50 = -3.2x - 0.3 \Rightarrow x = \frac{-49.7}{-3.2} = 15.531 \text{ 15.5 years from 1987 is 2002. According to our model, the trade deficit with China to reach $50 billion in 2002.}

d) Before. The last two points appear to be decreasing at a rate faster than the average rate of change.

e) The average yearly rate of change of the deficit is the slope of the line \( m = -3.2 \). Therefore, over the years considered, the deficit with China is increasing at an average of $3.2 \text{ billion/ year.}$

---

**To use a TI-83 or TI-84 to find the line of best fit**

- Enter your x and y values into lists. To do this, hit STAT and select 1: Edit... If you have anything in L1 and L2, cursor up to the name of the list, hit CLEAR and ENTER. Now just enter in your values one at a time by pressing ENTER after each number.
- To find the equation of the best fitting line, first hit STAT, cursor right to CALC and select option 4: LinReg(ax+b). Now press 2nd 1 , 2nd 2, to specify where you entered your data. (Don’t forget to put commas between the 2 variables). Now press ENTER. The equation of the line appears on your screen.
- To graph the best fit line with data points, first enter the equation of the line into Y1. Then make sure your stat plot is turned on. To do this, hit 2nd Y= and select 1...by pressing ENTER. You can then highlight any of the options you would like. Now press ZOOM and select 9: ZoomStat. You will have a plot of the original data points and the best fit line.