Using the TI-84+ to Graph, etc.

We will use the data from Ch.6 that shows the change in tuition costs from Arizona State University during the 90's.

Year (1990 = year 0):

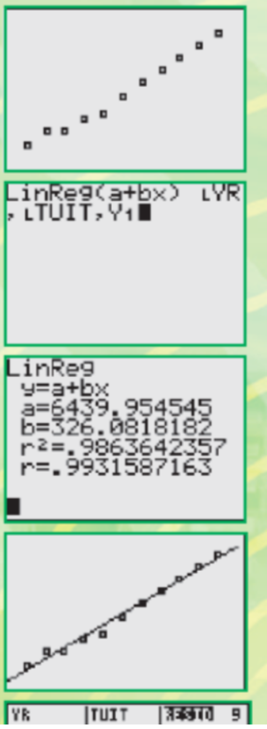
Pg. 13

0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10

Tuition ($):

6546, 6996, 6996, 7350, 7500, 7978,

8377, 8710, 9110, 9411, 9800



By now you will not be surprised to learn that your calculator can do it all: scatterplot, regression line, and residuals plot. Let's try it using the Arizona State tuition data from the last chapter. (TI Tips page 142) You should still have those saved in lists named YR and TUIT. First, recreate the scatterplot.

1. Find the equation of the regression line.

Actually, you already found the line when you used the calculator to get the correlation. But this time we'll go a step further so that we can display the line on our scatterplot. We want to tell the calculator to do the regression and save the equation of the model as a graphing variable.

- Under `STAT CALC` choose `LinReg(a+bx)`.
- Specify that `Xlist` and `Ylist` are `YR` and `TUIT`, as before, but . . .
- Now add one more specification to store the regression equation. Press `VARS`, go to the `Y-VARS` menu, choose `1:Function`, and finally(!) choose `Y1`.
- To `Calculate`, hit `ENTER`.

There's the equation. The calculator tells you that the regression line is $\widehat{tuit} = 6440 + 326\,year$. Can you explain what the slope and intercept mean?

2. Add the line to the plot.

When you entered this command, the calculator automatically saved the equation as `Y1`. Just hit `GRAPH` to see the line drawn across your scatterplot.

Residuals

- The model won't be perfect, regardless of the line we draw.
- Some points will be above the line and some will be below.
- The estimate made from a model is the predicted value (denoted as $\hat{y}$).
- The difference between the observed (true) value and the line's predicted value is called the residual.
- To find the residuals, we always subtract the predicted value from the observed one:

$$residual = \underbrace{observed}_{\text{will be given}} - \underbrace{predicted}_{\text{calc. from model}} = y - \hat{y}$$

The Least Squares Line Reminder:

- Since regression and correlation are closely related, we need to check the same conditions for regressions as we did for correlations:
  - Quantitative Variables Condition
  - Straight Enough Condition
  - Outlier Condition

Residuals Revisited

- The linear model assumes that the relationship between the two variables is a perfect straight line. The residuals are the part of the data that *hasn't* been modeled.

Data = Model + Residual

or (equivalently)

Residual = Data – Model

Or, in symbols,

$$e = y - \hat{y}$$

Residual = Actual - Predicted

 Chapter 7, Slide 19

$$e = \overset{\text{Actual}}{y} - \overset{\text{Predicted}}{\hat{y}}$$

A positive residual means:

* the actual observed value is greater than the model predicted value

* the model underestimated the actual value

Pts. above best fit line

A negative residual means:

* the actual observed value is less than the model predicted value

* the model overestimated the actual value

Pts. below best fit line

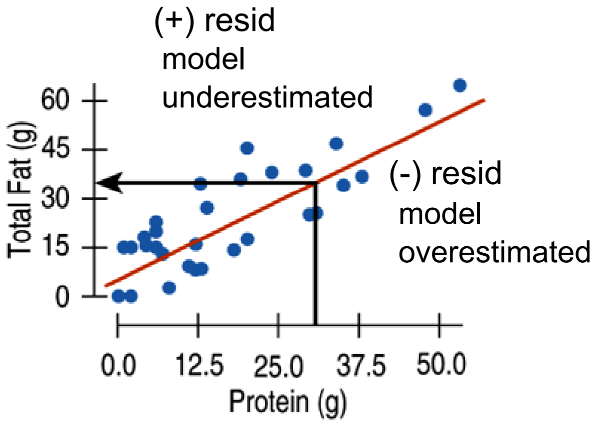
Residuals Revisited (cont.)

- Residuals help us to see whether the model makes sense.
- When a regression model is appropriate, nothing interesting should be left behind.
- After we fit a regression model, we usually plot the residuals in the hope of finding…nothing.

Fat Versus Protein: An Example

- The regression line for the Burger King data fits the data well:
  - The equation is

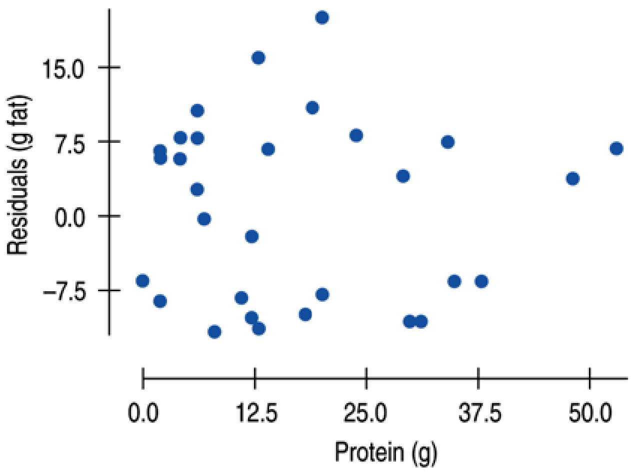
$$\widehat{fat} = 6.8 + 0.97\,protein.$$



The *predicted fat* content for a BK Broiler chicken sandwich (with 30 g of protein) is 6.8 + 0.97(30) = 35.9 grams of fat.

Residuals Revisited (cont.)

- The residuals for the BK menu regression look appropriately boring:



 Chapter 7, Slide 21

* Residual graphs should show NOTHING!

(meaning no patterns with especially high residuals grouped together.)

We want to see randomness in our Residual scatterplot.

* We can also check linear model appropriateness by looking at the scatterplot with the line graphed on top of the data. and r - value

Example: Look at Tuiton data with the line.

Just Checking

Resid = Actual - Predicted

Our linear model for Saratoga homes uses the *Size* (in thousands of square feet) to estimate the *Price* (in thousands of dollars): $\widehat{Price} = -3.117 + 94.45Size$. Suppose you're thinking of buying a home there.

13. Would you prefer to find a home with a negative or a positive residual? Explain.
14. You plan to look for a home of about 3000 square feet. How much should you expect to have to pay? size = 3
15. You find a nice home that size selling for $300,000. What's the residual?

Resid = actual - predicted
= 300,000 - 280,233
= 19,767

13. Negative; that indicates it's priced lower than a typical home of its size.
14. $280,233
15. $19,767 (positive!)

$\widehat{Price} = -3.117 + 94.45(3) = 280.233$

Your linear model for the class heights and weights was $\widehat{Weight} = -149 + 4.4Height$.

16. A student is 70" tall.
    a) Estimate her weight.
    b) She actually weighs 144 pounds. What's the residual?

17. Another student is 65" tall and has a residual of 18 pounds. How much does he weigh?

16. a) $-149 + 4.4(70) = 159$ pounds
    b) $144 - 159 = -15$ pounds

17. Estimate $-149 + 4.4(65) = 137$ pounds; 18 pounds heavier would be 155 pounds.

Resid = actual − predict

$$\begin{array}{r} 18 \\ +137 \\ \hline 155 \end{array} = \text{actual} - 137 \quad {+137}$$

Homework:

Textbook pg. 186-187

4, 6, 8, 10, 12, 14