Regression

Using regression to get a function that models the data.

Scatter plot: First we need to know how to use the calculator to draw a scatter plot. This will allow you to visually check the general shape of the data.

Step 1. The data must be entered into the calculator. Press [*STAT*] [ENTER]. This puts you in the edit menu. Enter the data into lists x $=L_1$ and y $=L_2$. If you choose any other list, just make the appropriate changes below. If some or all of the list are gone, you can get them back by pressing [*STAT*] [5] [ENTER]. To clear out a list, arrow up to the top of the list (highlighting its name) and press [CLEAR] [ENTER].

Step 2. This setup needs to be done only once. To set up the stat plots press [2nd] [Y=]. Choose your plot and make sure that it is on, x-list $=L_1$, y-list$= L_2$, and the type is the first graph on the first line. Once this is set up you can turn it on and off from the top of the [Y=] screen.

Step 3. Press [ZOOM] [9] (ZoomStat) to graph the scatter plot.

Lines: Given a table of values, we know that the data is linear if a constant change in x produces a constant change in y. Not all linear data sets meet this strict requirement. Data Set 1 does correspond to a linear function; however, it is not easy to notice this from the data. The scatter plot tells a different story. Data Set 2 has a constant change in x but not a constant change in the y. This means that it is not strictly linear. But looking at the scatter plot of this data shows its linear nature.

Data Set 1

x	2	4.5	8	9.2	12
y	16	23.5	34	37.6	46

Data Set 2

x	3	4	5	6
y	105	117	141	152

Now getting an equation for Data Set 1 is easy since we discovered (by the scatter plot) that it is linear. Pick any two points and compute the equation. Data Set 2 presents a different problem. It does have a linear nature, so which two points do we pick? Answer: Let the calculator do the job for us.

Having the calculator find the best fitting line to a set of data is called linear regression. Of course, this only works well if the data has a linear nature. To perform the regression, enter the data into the lists: x in $L_1$and y in L_2. To get the calculator to do the regression press [*STAT*] [▶] [4]. On the screen of the calculator you should see **LinReg(ax+b)**. If you used L_3 and L_4 for the data, now type [2nd] [3] [,] [2nd] [4]. On the screen you should now see **LinReg(ax+b) L_3,L_4**, now press enter.

Exponential: Given a table of values, our book tells us that they correspond to an exponential function if: the x-values have a difference of 1 and the ratios of the y values (a y-value divided by the previous y-value) is constant. Data Set 3 meets this condition. Notice $\frac{24}{16} = \frac{36}{24} = \frac{54}{36} = 1.5$ This means that in the formula $y = P_o a^x$, $a = 1.5$. Data Set 4 also corresponds to an exponential function. Data Set 5 seems to be exponential in nature. This is confirmed by the scatter plots for the different data sets.

Data Set 3

x	0	1	2	3
y	16	24	36	54

Data Set 4

x	1	4	5	7
y	3	24	48	192

Data Set 5

x	1	3	4	6	7
y	23.5	70.3	169.8	381.3	703.7

The calculator does exponential regression also. Press [*STAT*] [▶] [0]. On the screen of the calculator you should see **ExpReg**. If you used L_3 and L_4 for the data, now type [2nd] [3] [,] [2nd] [4]. On the screen you should see **ExpReg L_3,L_4**, now press enter.