Chapter 4: More About Bivariate Relationships

4.1 Transforming to Achieve Linearity

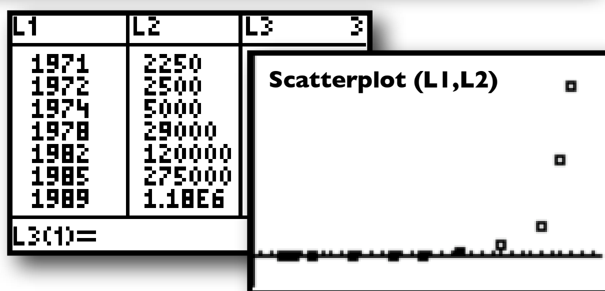
In chapter 4, we learn how to find nonlinear models for data by transforming to achieve linearity. The transformation process can be very time consuming, however, we can use our calculator to automate some of our work. It is unlikely you'd be asked to find a power or exponential model on the AP exam. However, knowing the procedure will lead to a better understanding of regression models. *Note: Like the previous chapters, be sure you understand the concepts behind these methods before using the calculator!*

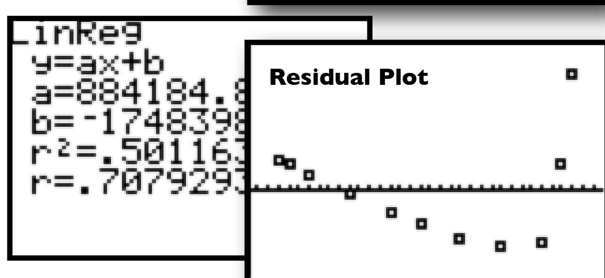
Modeling Exponential Growth

Consider the following data on the number of transistors on a chip. According to "Moore's Law," the number grows exponentially over time. {In 1965, Gordon Moore, a co-founder of Intel, actually predicted the number would double every 18 months.}

Year	1971	1972	1974	1978	1982	1985	1989	1993	1997	1999	2000
Transistors	2250	2500	5000	29000	120000	275000	1180000	3100000	7500000	24000000	42000000

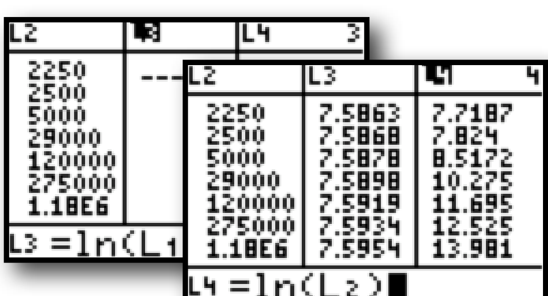
1. Enter the data into **L1** and **L2**.
2. **ZOOM STAT** to a Scatterplot
3. Calculate the LSRL
4. View the Residual Plot
5. Press ENTER

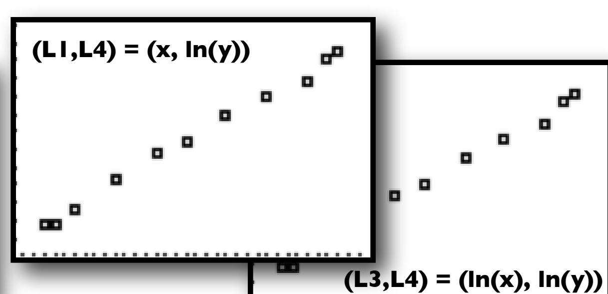




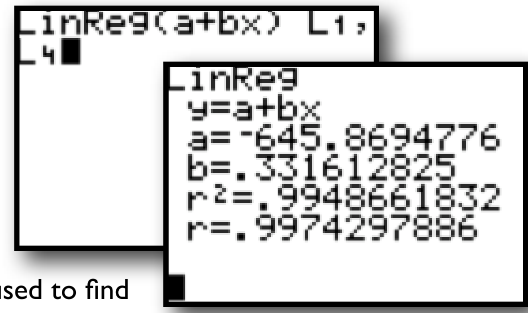
Note the relatively low coefficient of correlation and the distinct pattern in both the scatterplot and residual plot. We have evidence to suggest a nonlinear model may be more appropriate.

To find the "best fitting" nonlinear model, we must transform our data to achieve linearity. To do this, define **L3** to be ln(x) and **L4** to be ln(y). As we learned in Chapter 4, if (x, ln(y)) is linear, then an exponential model may be appropriate. If (ln(x),ln(y)) is linear, a power model may be appropriate. Use the list editor to transform your data. Then, **ZOOM STAT** to plots of (**L1, L4**) and (**L3,L4**).





Notice, both (x, ln(y)) and (ln(x), ln(y)) are approximately linear. It appears either an exponential or a power model could be used. Since the situation appears to be exponential in nature, we will proceed to build an exponential model.



1. Find the LSRL of (x, ln(y)).
2. Transform (ln(y))-hat = a+bx back to the original (x,y) context.
3. y-hat = $e^{(a+bx)}$
4. *transistor-hat* = $e^{(-645.87+.331(year))}$

Note: **STAT CALC ExpReg L1,L2** can also be used to find this model, but may result in an "overflow" error.

Modeling Power Growth

The process for modeling power growth is the same as that for modeling exponential growth.

1. Enter your data and check the scatterplot.
2. If linear, find the LSRL. If not linear...
3. Transform your data by defining **L3=ln(L1) and L4=ln(L2).**
4. Check both **(L1, L4)** and **(L3, L4)** scatterplots for linearity. Consider *r* and residual plots if necessary. Linear **(L1, L4)** implies exponential. Linear **(L3,L4)** implies power model.
5. Find the LSRL of the 'more linear' transformed data.
6. Transform the Linear Regression Equation back to the original data.

If **(L1, L4)** was more linear, ln(y)-hat = a + bx

Exponential Model: y-hat = $e^{(a+bx)}$

If **(L3, L4)** was more linear, ln(y)-hat = a + b(ln(x))

Power Model: y-hat = $e^a x^b$

While the TI is capable of finding an exponential or power model through STAT CALC, it helps to know *how* it goes about doing so. Again, the chances are minimal that you'd have to go through this entire procedure on the AP Exam. Make sure you understand the process as you may be asked to describe it or justify the use of a particular model.

AP® Examination Tips

When taking the Advanced Placement Statistics Exam, the chances are slim that you would need to perform a transformation to achieve linearity. You may be asked to interpret the results of such a transformation, or to justify the use of a particular regression model based on graphical and numeric output. When doing so, be sure to interpret the output in the context of the problem!

When finding a non-linear model

- Start with an analysis of the original (x,y) data
- Check BOTH (x, ln(y)) and (ln(x), ln(y))
- Refer to *r* and residual plots when determining which is "more linear"
- Be sure to keep your lists straight...don't forget which is ln(x) or ln(y)
- LABEL sketches with appropriate variables!

When interpreting results

- Keep your variables straight...are you referring to y or ln(y)?
- Remember, linear (x, ln(y)) implies exponential while linear (ln(x), ln(y)) implies a power model may be more appropriate.
- Be careful when using models to "plug in" appropriate measures and to interpret predicted values correctly.
- Be careful not to extrapolate too far beyond the observed data.