**Section 12.2 - Transforming to Achieve Linearity Part 2** (pp. 771-785)

**Transforming with logarithms** - Not all curved relationships are described by power models. Some relationships can be described by a logarithmic model of the form

 y = a + b log x.

Sometimes the relationship between y and x is based on repeated multiplication by a constant factor. That is, each time x increases by 1 unit, the value of y is multiplied by b. An exponential model of the form y = abx describes such multiplicative growth.

If an exponential model of the form y = abx describes the relationship between x and y, we can use logarithms to transform the data to produce a linear relationship.

We can rearrange the final equation as log y = log a + (log b)x. Notice that log a and log b are constants because a and b are constants.

 So the equation gives a linear model relating the explanatory variable x to the transformed variable
log y.

Thus, if the relationship between two variables follows an exponential model, and we plot the logarithm (base 10 or base e) of y against x, we should observe a straight-line pattern in the transformed data.

If we fit a least-squares regression line to the transformed data, we can find the predicted value of the logarithm of y for any value of the explanatory variable x by substituting our x-value into the equation of the line.

 To obtain the corresponding prediction for the response variable y, we have to “undo” the logarithm transformation to return to the original units of measurement. One way of doing this is to use the definition of a logarithm as an exponent:

**Example** - Moore’s Law and Computer Chips (p. 773)

Gordon Moore, one of the founders of Intel Corporation, predicted in 1965 that the number of transistors on an integrated circuit chip would double every 18 months. This is Moore’s law, one way to measure the revolution in computing. Here are data on the dates and number of transistors for Intel microprocessors:

|  |  |
| --- | --- |
| Screen shot 2010-12-06 at 1.41.33 PM.png | Screen shot 2010-12-06 at 1.41.43 PM.png |

(a) A scatterplot of the natural logarithm (log base e or ln) of the number of transistors on a computer chip versus years since 1970 is shown. Based on this graph, explain why it would be reasonable to use an exponential model to describe the relationship between number of transistors and years since 1970.

(b) Minitab output from a linear regression analysis on the transformed data is shown below. Give the equation of the least-squares regression line. Be sure to define any variables you use.

|  |  |
| --- | --- |
| Screen shot 2010-12-06 at 1.43.28 PM.png | Screen shot 2010-12-06 at 1.43.40 PM.png |

(d) A residual plot for the linear regression in part (b) is shown below. Discuss what this graph tells you about the appropriateness of the model.



**Power Models Again**

When we apply the logarithm transformation to the response variable y in an exponential model, we produce a linear relationship. To achieve linearity from a power model, we apply the logarithm transformation to both variables. Here are the details:

A power model has the form y = axp, where a and p are constants.

Take the logarithm of both sides of this equation. Using properties of logarithms,

 log y = log(axp) = log a + log(xp) = log a + p log x

 The equation log y = log a + p log x shows that taking the logarithm of both variables results in a linear relationship between log x and log y.

3. Look carefully: the power p in the power model becomes the slope of the straight line that links log y to log x.

*If a power model describes the relationship between two variables, a scatterplot of the logarithms of both variables should produce a linear pattern. Then we can fit a least-squares regression line to the transformed data and use the linear model to make predictions.*

**Example** (p. 778) - On July 31, 2005, a team of astronomers announced that they had discovered what appeared to be a new planet in our solar system. Originally named UB313, the potential planet is bigger than Pluto and has an average distance of about 9.5 billion miles from the sun. Could this new astronomical body, now called Eris, be a new planet? At the time of the discovery, there were nine known planets in our solar system. Here are data on the distance from the sun (in astronomical units, AU) and period of revolution of those planets.

|  |  |
| --- | --- |
| Screen shot 2010-12-06 at 2.23.05 PM.png | Screen shot 2010-12-06 at 2.22.38 PM.png |

Describe the relationship between distance from the sun and period of revolution.

(a) Based on the scatterplots below, explain why a power model would provide a more appropriate description of the relationship between period of revolution and distance from the sun than an exponential model.



(b) Minitab output from a linear regression analysis on the transformed data (ln(distance), ln(period)) is shown below. Give the equation of the least-squares regression line. Be sure to define any variables you use.



(c) Use your model from part (b) to predict the period of revolution for Eris, which is 9,500,000,000/93,000,000 = 102.15 AU from the sun. Show your work.

(d) A residual plot for the linear regression in part (b) is shown below. Do you expect your prediction in part (c) to be too high, too low, or just right? Justify your answer.



HW: p. 788 problems 37, 39, 41, 45-50.