**Section 12.2 - Transforming to Achieve Linearity Part 1** (pp. 765-771)

In Chapter 3, we learned how to analyze relationships between two quantitative variables that showed a linear pattern. When two-variable data show a curved relationship, we must develop new techniques for finding an appropriate model. This section describes several simple transformations of data that can straighten a nonlinear pattern.

Once the data have been transformed to achieve linearity, we can use least-squares regression to generate a useful model for making predictions. And if the conditions for regression inference are met, we can estimate or test a claim about the slope of the population (true) regression line using the transformed data.

Applying a function such as the logarithm or square root to a quantitative variable is called transforming the data. We will see in this section that understanding how simple functions work helps us choose and use transformations to straighten nonlinear patterns.

**Example** - Imagine that you have been put in charge of organizing a fishing tournament in which prizes will be given for the heaviest Atlantic Ocean rockfish caught. You know that many of the fish caught during the tournament will be measured and released. You are also aware that using delicate scales to try to weigh a fish that is flopping around in a moving boat will probably not yield very accurate results. It would be much easier to measure the length of the fish while on the boat. What you need is a way to convert the length of the fish to its weight.



|  |  |
| --- | --- |
| F12.12.jpgScatterplot of rockfish weight vs. length | F12.13.jpgScatterplot of weight vs. length3 |



 Scatterplot of $\sqrt[3]{weight}$ vs. length

Here is Minitab output from separate regression analyses of the two sets of transformed Atlantic Ocean rockfish data.

|  |
| --- |
| tableun_12_15.jpg |
| F12.UN34.jpg |

When experience or theory suggests that the relationship between two variables is described by a power model of the form y = axp, you now have two strategies for transforming the data to achieve linearity.

1. Raise the values of the explanatory variable x to the p power and plot the points $(x^{p},y)$.

2. Take the pth root of the values of the response variable y and plot the points $(x, \sqrt[p]{y})$

What if you have no idea what power to choose? You could guess and test until you find a transformation that works. Some technology comes with built-in sliders that allow you to dynamically adjust the power and watch the scatterplot change shape as you do.

HW: p. 786 problems 33, 35.