### A4. Regression Models

#### Power Law Growth

The technique of transforming data, especially using logarithms, is extremely valuable. Many physical systems involve two variables \(X\) and \(Y\) that are known (or suspected) to obey a "power law" relation, where \(Y\) is proportional to \(X\) raised to a power, i.e., \(Y = \alpha X^\beta\) for some fixed constants \(\alpha\) and \(\beta\). Examples include the relation \(L = 1.4 A^{0.6}\) that exists between river length \(L\) and the area \(A\) that it drains, “inverse square” laws such as the gravitational attraction \(F = G m_1 m_2 r^{-2}\) between two masses separated by a distance \(r\), earthquake frequency versus intensity, the frequency of global mass extinction events over geologic time, comet brightness vs. distance to the sun, economic trends, language patterns, and numerous others.

As mentioned before, in these cases, *both* variables – \(X\) and \(Y\) – are often transformed by means of a logarithm. The resulting data are replotted on a “log-log” scale, where a linear model is then fit: (The algebraic details were presented in the basic review of logarithms.)

\[
\log_{10} Y = \beta_0 + \beta_1 \log_{10} X,
\]

and the original power law relation can be recovered via the formulas

\[\alpha = 10^{\hat{\beta}_0} \quad \beta = \hat{\beta}_1.\]

As a simple example, suppose we are examining the relation between \(V = \text{Volume (cm}^3)\) and \(A = \text{Surface Area (cm}^2)\) of various physical objects. For the sake of simplicity, let us confine our investigation to sample data of solid cubes of \(n = 10\) different sizes:

<table>
<thead>
<tr>
<th>(V)</th>
<th>1</th>
<th>8</th>
<th>27</th>
<th>64</th>
<th>125</th>
<th>216</th>
<th>343</th>
<th>512</th>
<th>729</th>
<th>1000</th>
</tr>
</thead>
<tbody>
<tr>
<td>(A)</td>
<td>6</td>
<td>24</td>
<td>54</td>
<td>96</td>
<td>150</td>
<td>216</td>
<td>294</td>
<td>384</td>
<td>486</td>
<td>600</td>
</tr>
</tbody>
</table>

Note the nonlinear scatterplot in Figure 1. If we take the common logarithm of both variables, the rescaled “log-log” plot reveals a strong *linear* correlation; see Figure 2. This is strong evidence that there is a power law relation between the original variables, i.e., \(A = \alpha V^\beta\).

\[
\begin{align*}
\log_{10} V & = 0.000  & 0.903  & 1.431  & 1.806  & 2.097  & 2.334  & 2.535  & 2.709  & 2.863  & 3.000 \\
\log_{10} A & = 0.778  & 1.380  & 1.732  & 1.982  & 2.176  & 2.334  & 2.468  & 2.584  & 2.687  & 2.778 \\
\end{align*}
\]

Therefore, a linear model will be a much better fit for these *transformed* data points than for the original data points. Solving for the regression coefficients in the usual way *(Exercise)*, we find that the least squares regression line is given by

\[
\log_{10} V = \hat{\beta}_0 + \hat{\beta}_1 \log_{10} A.
\]
We can now estimate the original coefficients: $\hat{\alpha} = 10^{0.778} = 6$, and $\hat{\beta} = 0.667 = \frac{2}{3}$, approximately. Therefore, the required power law relation is $A = 6V^{2/3}$. This should come as no surprise, because the surface area of a cube (which has six square faces) is given by $A = 6s^2$, and the volume is given by $V = s^3$, where $s$ is the length of one side of the cube. Hence, eliminating the $s$, we see that $A = 6V^{2/3}$ for solid cubes. If we had chosen to work with spheres instead, only the constant of proportionality $\alpha$ would have changed slightly (to $\frac{3}{3 \sqrt{36\pi}} \approx 4.836$); the power would remain unchanged at $\beta = \frac{2}{3}$. (Here, $V = \frac{4}{3}\pi r^3$ and $A = 4\pi r^2$, where $r$ is the radius.) This illustrates a basic principle of mechanics: since the volume of any object is roughly proportional to the cube of its “length” (say, $V \propto L^3$), and the surface area is proportional to its square (say, $A \propto L^2$), what follows is the general power relation that $A \propto V^{2/3}$.

Comment. In a biomechanical application of power law scaling, consider the relation between the metabolic rate $Y$ of organisms (as measured by the amount of surface area heat dissipation per unit time), and their body mass $M$ (generally proportional to the volume). From the preceding argument, one might naively expect that, as a general rule, $Y \propto M^{2/3}$. However, this has been shown not to be the case. From systematic measurements of the correlation between these two variables (first done in 1932 by Max Kleiber), it was shown that a more accurate power relation is given by $Y \propto M^{3/4}$, known as Kleiber’s Law. Since that time, “quarter-power scaling” has been shown to exist everywhere in biology, from respiratory rates ($\propto M^{-1/4}$), to tree trunk and human aorta diameters ($\propto M^{3/8}$). Exactly why this is so universal is something of a major mystery, but seems related to an area of mathematics known as “fractal geometry.” Since 1997, much research has been devoted to describe general models that explain the origin and prevalence of quarter-power scaling in nature, and is considered by some to be “perhaps the single most pervasive theme underlying all biological diversity.” (Santa Fe Institute Bulletin, Volume 12, No. 2.)
Figure 1

Figure 2