16.2: Log Transformations

Learning Objectives

- State how a log transformation can help make a relationship clear
- Describe the relationship between logs and the geometric mean

The log transformation can be used to make highly skewed distributions less skewed. This can be valuable both for making patterns in the data more interpretable and for helping to meet the assumptions of inferential statistics. Figure $16.2.1$ shows an example of how a log transformation can make patterns more visible. Both graphs plot the brain weight of animals as a function of their body weight. The raw weights are shown in the upper panel; the log-transformed weights are plotted in the lower panel.

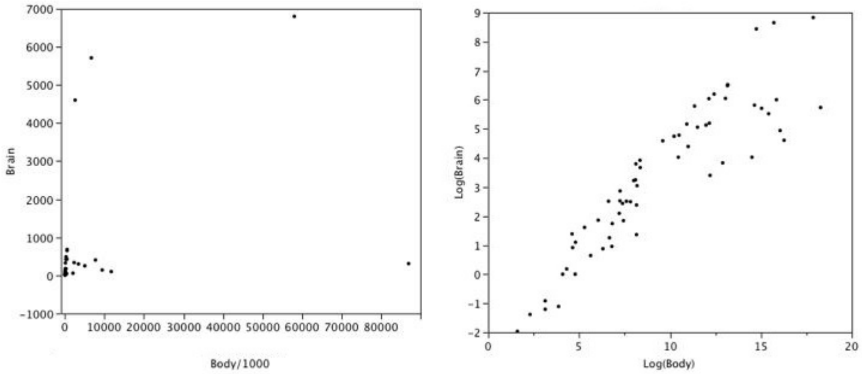

Figure 16.2.1: Scatter plots of brain weight as a function of body weight in terms of both raw data (upper panel) and log-transformed data (lower panel)

It is hard to discern a pattern in the upper panel whereas the strong relationship is shown clearly in the right panel. The comparison of the means of log-transformed data is actually a comparison of **geometric means**. This occurs because, as shown below, the anti-log of the arithmetic mean of log-transformed values is the geometric mean.

Table $16.2.1$ shows the logs (base 10) of the numbers 1, 10, and 100. The arithmetic mean of the three logs is

$$(0+1+2)/3=1 \tag{16.2.1}$$

The anti-log of this arithmetic mean of 1 is

$$10^1=10 \tag{16.2.2}$$

which is the geometric mean:

$$(1\times 10\times 100)^{0.3333}=10 \tag{16.2.3}$$

Table 16.2.1: Logarithms

X	$\log_{10}(X)$
1	0
10	2
100	3
1,000	4
10,000	5

Therefore, if the arithmetic means of two sets of log-transformed data are equal, then the geometric means are equal.