Figure 13.7. Lognormal distribution of species abundance. As sampling intensity increases, the veil line moves to the left, revealing more of the relative abundance distribution. The area under the exposed portion of the curve on the right is the observed species richness. The area under the entire curve is the estimated total species richness of the community. Estimating species richness using this method requires large data sets that clearly reveal a mode (the highest point of the curve) and closely fit a lognormal distribution.

log scale. In such a plot the lognormal, as the name implies, forms a normal distribution. Preston (1948) proposed that if a community is undersampled, only the rightmost part of the curve is revealed. He coined the term *veil line* for a vertical line dividing the lognormal distribution into two portions, the rightmost portion being the more abundant species revealed by sampling and the leftmost portion being the less abundant species remaining to be sampled. He proposed that as sampling increases the veil line moves to the left, revealing more and more of the curve. Thus sample data may be interpreted as a truncated lognormal distribution (Pielou 1975).

In practice, only very large data sets have revealed lognormal distributions. If the mode of a lognormal curve (the highest point in the distribution) is not revealed, it is practically impossible to distinguish a truncated lognormal from a log-series (Magurran 1988). Lambshead and Platt (1985) argue that the shape of the lognormal distribution should be independent of sample size and that there is no evidence that the veil line moves to the left as sample size

increases. Hughes (1986) even suggests that some of the observed lognormal distributions could be caused by species misidentifications and sampling errors.

In constructing frequency histograms of species abundances, the traditional practice is to use $\log_2$ for the horizontal axis, so that each abundance class represents a doubling of the previous one. In constructing observed distributions from real data, abundance classes are defined and the number of species in each abundance class tallied. Ideally abundance is a continuous variable such as biomass or cover, but typically abundance is number of individuals. Problems arise when fitting discontinuous abundance data to a continuous distribution such as the lognormal. Different methods have been proposed for defining abundance classes. Ludwig and Reynolds (1988) describe a common way, which is to define abundance classes as 0–1 individuals, 1–2, 2–4, 4–8, 8–16, and so on. For each species that straddles abundance classes (i.e., with abundance 1, 2, 4, 8, 16, and so on), 0.5 is added to the tally for each of the adjacent abundance classes. A problem with this method is that singletons are split between two abundance classes. The lowest abundance class will contain half the singletons, and the second lowest abundance class will contain half the singletons plus half the doubletons. This method forces the second lowest abundance class to have more species than the lowest in all cases, and it thus gives the false impression that the mode of a lognormal distribution has been revealed (Colwell and Coddington 1994).

Magurran uses an alternative method of defining abundance classes. She defines the lowest abundance class as the sum of all the singletons and doubletons, the next lowest as species with abundance 3 or 4, the next 5–8, then 9–16, 17–32, and so on. This method does not generate a "pseudo-mode" at the second abundance class. The Berlese data set shows that the same data can appear radically different