

OLM12.1 Comparison of species numbers in samples

In the discussion of diversity patterns in Ch12Divers we have omitted the difficulties stemming in the fact that it is just samples whose species numbers can be compared in diversity studies, not the entire communities. Just as the finite number of individuals in an offspring generation do not necessarily inherit all the alleles present in the parent population (TBox 11.4, p.243), the species of a community are usually not all represented in the samples (Figure 12.1.1). This OLM deals with the methods of making comparative estimates of diversity even in the face of such difficulties.

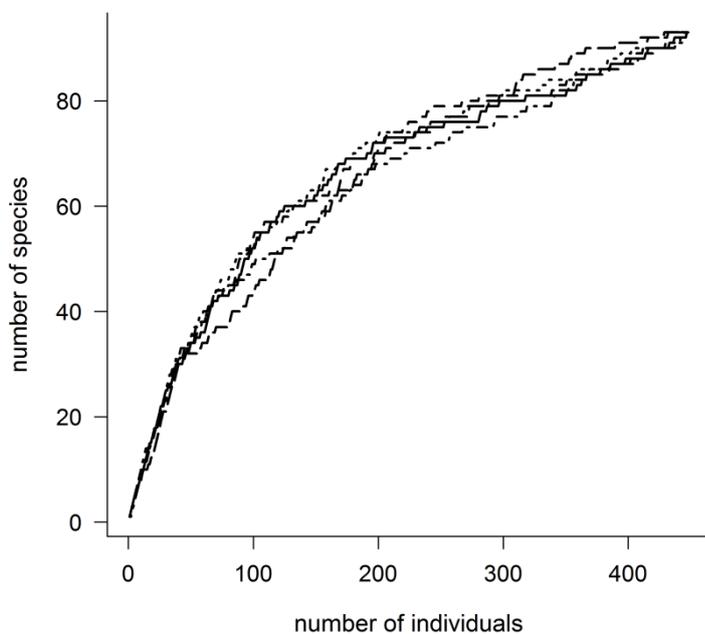


Figure 12.1.1 Uncertainty of the number of species in a sample

Having counted all the trees with diameter at breast height (DBH) over 10 cm within a 1 ha area of Baro Colorado Island, the number of individuals was found to be 448, distributed in 93 species (Condit et al. 2002). The figure shows how the cumulative number of species increases, if the individuals are encountered at 5 different random orders. Each of the curves necessarily starts at 1 individual – 1 species, and ends at 448 individuals – 93 species, but the actual course of the graph depends on the order of inclusion of the individuals in the sample (data from Condit et al. 2002).

Since the number of species within the sample cannot exceed that of the community, the expected value of the species numbers within the samples is always an underestimation of the actual number. The bias decreases with the size of the sample, but in very diverse communities and in ones with a high number of rare species the sample size sufficient to reduce the bias to a negligible level may be very large. If a randomly chosen individual belongs to a certain species with probability p , then the chance of that species to be included in a random sample of size N is $1-(1-p)^N$. If the species is rare (p is small), then this chance comes close to 1 only at huge sample sizes (Figure 12.1.2).

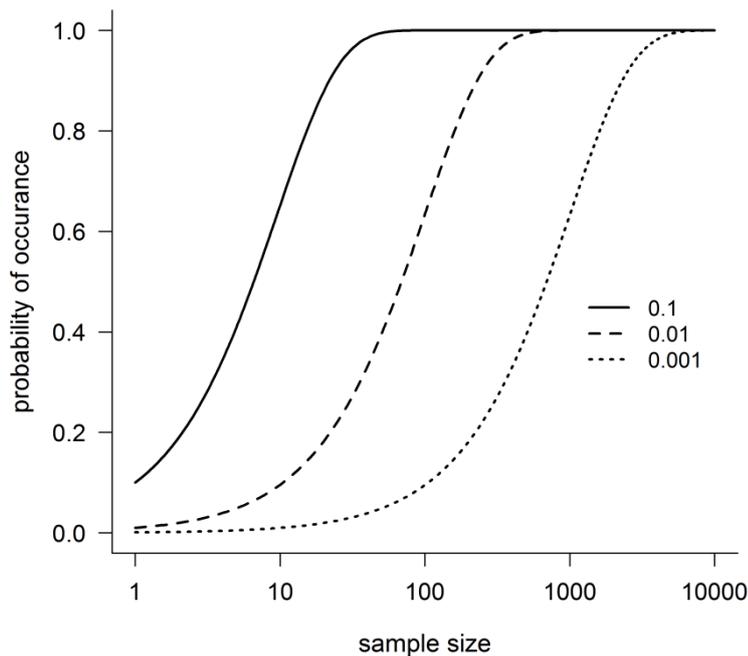


Figure 12.1.2 Effect of the degree of rarity on detection probability.

The probability of occurrence of three species with three different frequencies in samples of increasing sizes, assuming that the individuals are included in the samples independently. Under this assumption, choosing N individuals into the sample, the chance of a species of relative frequency p to avoid being chosen at all is $(1-p)^N$; therefore, the chance of being included in the sample at least once is that of the complementary event: $1-(1-p)^N$.

Erwin (1982, 1991) has collected and determined 1100 beetle species from the canopy of 19 individuals of a tropical tree species, *Luehea seemannii*, showing that even 1000 species does not represent an extreme species diversity in tropical insect communities. From such a community of 1000 evenly distributed species the sample size sufficient for an expected 999 species per sample requires a sample size of 6900 individuals. This is an over-optimistic estimate, because if the species are not equally frequent, the expected number of species per sample decreases. If, for example, the number of species is again 1000 within the community, but they are distributed in 9 frequent species with 10% of the individuals in each and the rest is distributed evenly among rare species, then the sample size sufficient to include 999 species on average will be ten times larger.

A possible solution to the underestimation problem would be to estimate the bias as well, and to give an unbiased estimation for the species diversity of the community on that basis. We suggest that the Reader interested in such methods turn to Colwell and Coddington (1994), O'Hara (2005) and Chao et al. (2014) for details.

Another possible solution is to compare samples of equal sizes, i.e., including the same number of individuals, or resulting from the same sampling effort (e.g., using the same number of traps for the same length of time, or – in case of plants – sampling areas of the same extension). These two methods may give contradictory results Figure 12.1.3 so it is advisable to make a clear difference of them in any actual case. The result should be called *species richness* the first case and *species density* in the second, but in many cases both are referred to as species diversity.

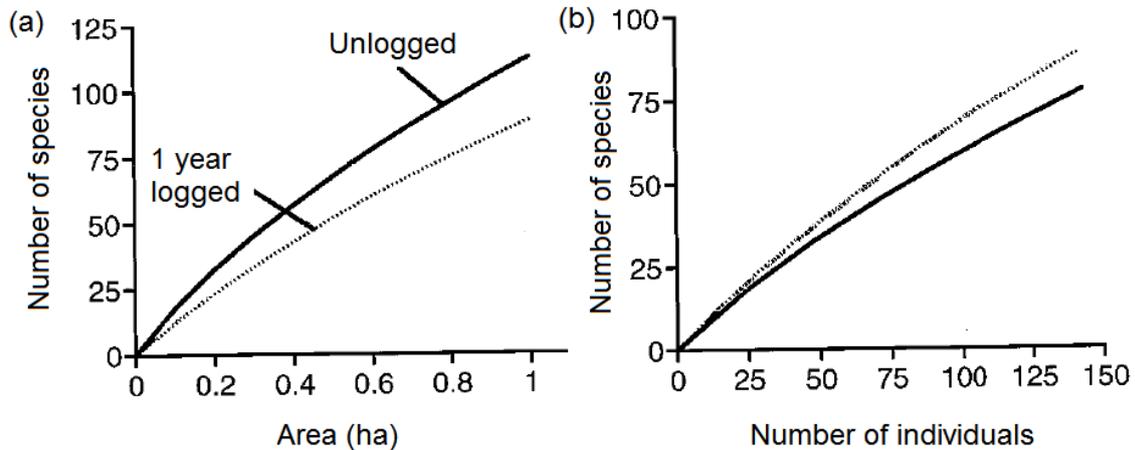


Figure 12.1.3 Cumulative number of species in a selectively logged tropical rainforest in Indonesian Borneo. a) Cumulative species-area relationships and b) cumulative species-individual relationship for all trees >20 cm in diameter. In the case of area-based sampling (panel a) the unlogged stand seems more diverse, but with fixed numbers of individuals per sample (panel b) it is the logged stand that shows higher species numbers (From Cannon et al. 1998). Reprinted with permission from AAAS.

Comparing the species diversity of samples of different sizes is possible by calculating the expected number of species in samples of equal size n , where n is not larger than the smallest of the samples to be compared. This is called the method of *rarefaction*, which means re-sampling the samples without replacement to be compared, by choosing n individuals from each, and calculating the average number of species in the sub-samples thus obtained.

References

- Cannon, C.H., Peart, D.R. and Leighton, M. (1998). Tree Species Diversity in Commercially Logged Bornean Rainforest. *Science*, 281(5381): 1366-68.
- Chao, A., Gotelli, N.J., Hsieh, T.C., et al. (2014). Rarefaction and extrapolation with Hill numbers: a framework for sampling and estimation in species diversity studies. *Ecological Monographs*, 84(1): 45-67.
- Colwell, R.K. and Coddington, J.A. (1994). Estimating terrestrial biodiversity through extrapolation. *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences*, 345(1311): 101-18.
- Condit, R., Pitman, N., Leigh, E.G., et al. (2002). Beta-diversity in tropical forest trees. *Science*, 295(5555): 666-69.
- Erwin, T.L. (1982). Tropical Forests: Their Richness in Coleoptera and Other Arthropod Species. *The Coleopterists Bulletin*, 36(1): 74-75.
- Erwin, T.L. (1991). How Many Species Are There?: Revisited. *Conservation Biology*, 5(3): 330-33.
- O'Hara, R.B. (2005). Species richness estimators: how many species can dance on the head of a pin? *Journal of Animal Ecology*, 74(2): 375-86.