Coverage-Based Rarefaction and Extrapolation:

Standardizing Samples by Completeness Rather Than Size

趙蓮菊 清華大學統計學研究所

Abstract

I will talk on an integrated sampling, rarefaction, and extrapolation methodology to compare species richness of a set of communities based on samples of equal completeness (as measured by sample coverage) instead of equal size. Traditional rarefaction or extrapolation to equal-sized samples can misrepresent the relationships between the richnesses of the communities being compared, because a sample of a given size may be sufficient to fully characterize the lower-diversity community but insufficient to characterize the richer community. Thus the traditional method systematically biases the degree of differences between community richnesses. We derive a new analytic method for seamless coverage-based rarefaction and extrapolation. We show that this method yields less-biased comparisons of richness between communities, and manages this with less total sampling effort. When this approach is integrated with an adaptive coverage-based stopping rule during sampling, samples may be compared directly without rarefaction, so no extra data is taken and none is thrown away. Even if this stopping rule is not used during data collection, coverage-based rarefaction throws away less data than traditional size-based rarefaction, and more efficiently finds the correct ranking of communities according to their true richnesses. Several hypothetical and real examples demonstrate these advantages.