Improving Species Richness Estimation under Heterogeneous Detection and Integrated Sampling Data

Chun-Huo Chiu (邱春火)
Department of Agronomy, National Taiwan University, Taiwan
chchiu2017@ntu.edu.tw

Abstract

Species richness is one of the most intuitive and widely used measures of biodiversity. In practice, complete detection of all species in a community is rarely achievable, and observed richness is therefore systematically underestimated. This problem is particularly severe in highly heterogeneous communities or when sample sizes are small. Numerous estimators have been proposed to correct this bias, and these methods are generally classified as parametric or nonparametric. Among nonparametric approaches, Chao's lower-bound estimator is the most widely applied because it requires minimal assumptions and performs robustly under heterogeneous detection probabilities.

Despite these advantages, Chao's estimator has two important limitations. It can seriously underestimate richness in extremely heterogeneous assemblages, and it cannot be directly applied to integrated data that combine species abundance and incidence information. In this study, I propose a new framework to address both issues. Parametric mixture models are used to reduce the downward bias of Chao's lower bound, and sampling distribution theory is employed to extend the estimator to integrated data. Simulation results show that the proposed estimator substantially reduces bias, exhibits more robust statistical behavior than existing parametric estimators, and provides more accurate confidence interval coverage, especially in highly heterogeneous communities. An application to a real ecological dataset further demonstrates the practical usefulness of the proposed approach.