

**A Computationally Efficient Approach to Estimating
Species Richness and Rarefaction Curve**

Seungchul Baek, PhD
Assistant Professor
Department of Mathematics & Statistics
University of Maryland, Baltimore County

12:00 Noon Eastern time, Tuesday, October 26, 2021

ABSTRACT

In ecological and educational studies, estimators of the total number of species and rarefaction curve based on empirical samples are important tools. We propose a new method to estimate both rarefaction curve and the number of species based on a ready-made numerical approach such as quadratic optimization. The key idea in developing the proposed algorithm is based on nonparametric empirical Bayes estimation incorporating an interpolated rarefaction curve through quadratic optimization with linear constraints based on g -modeling in Efron (2014). Our proposed algorithm is easily implemented and shows better performances than existing methods in terms of computational speed and accuracy. Furthermore, we provide a criterion of model selection to choose some tuning parameters in estimation procedure and the idea of confidence interval based on asymptotic theory rather than resampling method. We present some asymptotic result of our estimator to validate the efficiency of our estimator theoretically. A broad range of numerical studies including simulations and real data examples are also conducted, and the gain that it produces has been compared to existing methods.