

Finding happiness in your SAD

Andrew J. Rominger

Outline - Why binning is bad - more error in estimation (e.g. estimate SAD on raw v. binned data) - modes mean little - gambin is bastardized negbinom - Issues with sampling - parameters scale with sample size - parametric fit is still pretty good - this is different from the veil line which is a fraught concept - combining samples conflates poisson with negbinom sampling - The core debate is between lnorm and gamma and that might not be winnable with finite data - most models come from one or the other - likelihood equivilance across param space of the two models - What do real world SADs look like - which models win - for winning model, where in param space do they fall

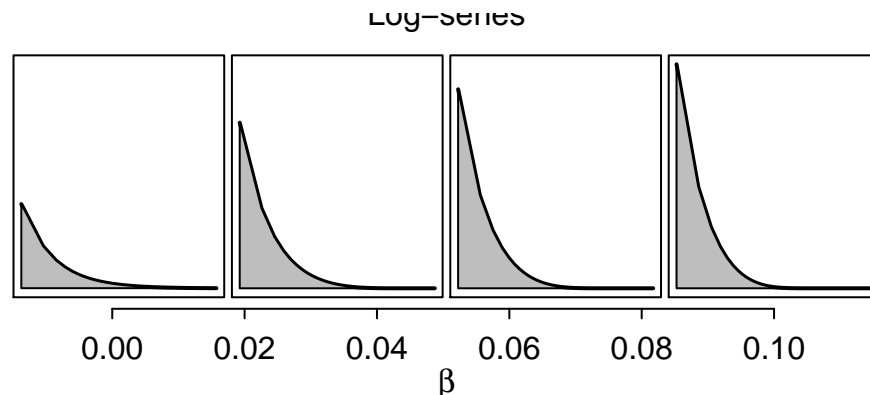
Code to-do - $P(S | N)$ and $P(N | S)$ - $P(n_i | S, N)$ - modes - model fit with $P(n_i | S, N)$ - param scaling under sampling - lnorm gamma param space equililance - real world SADs - model fit - param space density

Abstract

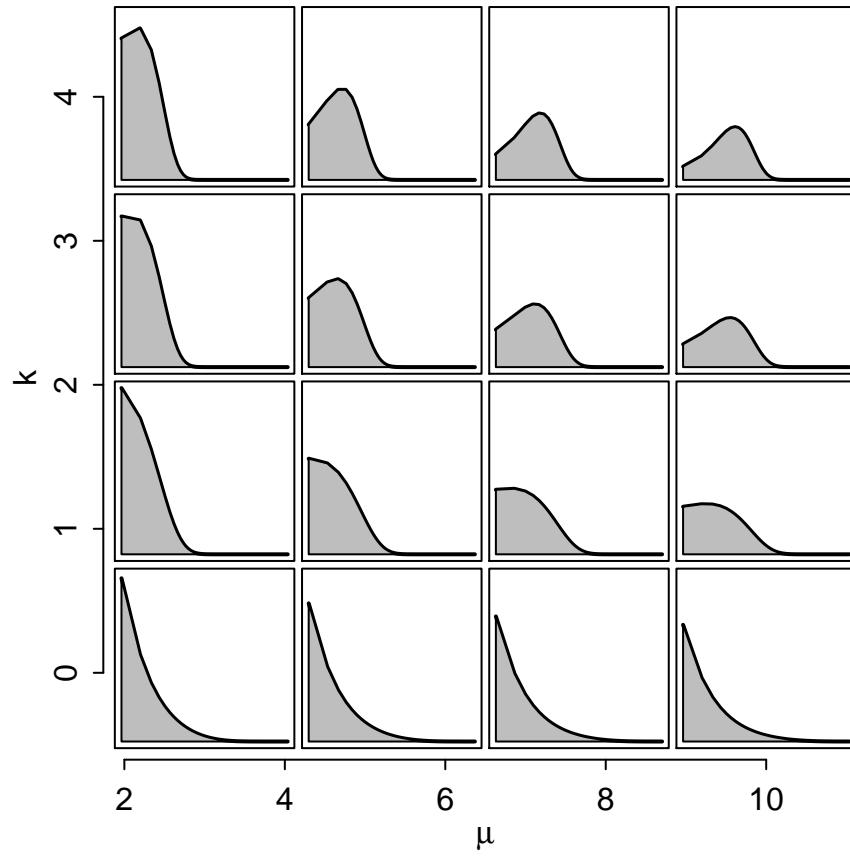
Introduction

The species abundance distribution (SAD) is a central metric used throught Ecology to describe the commonness

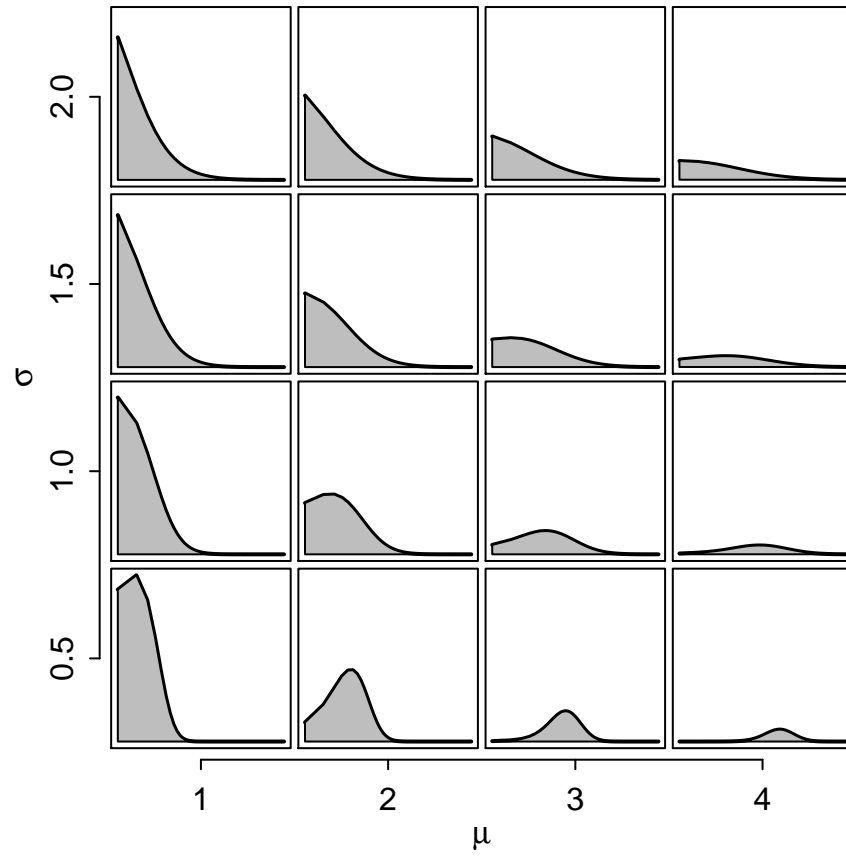
SAD shapes



Truncated negative binomial



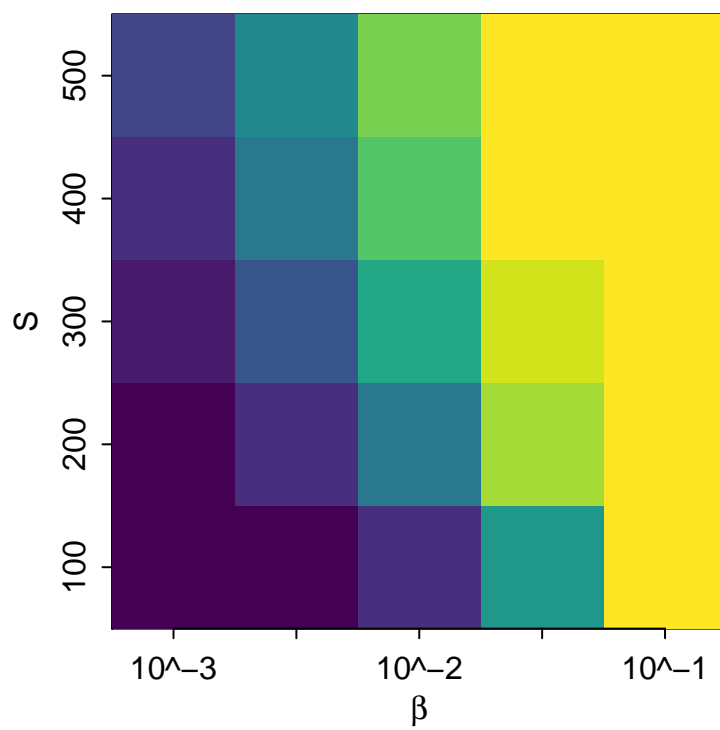
Poisson lognormal



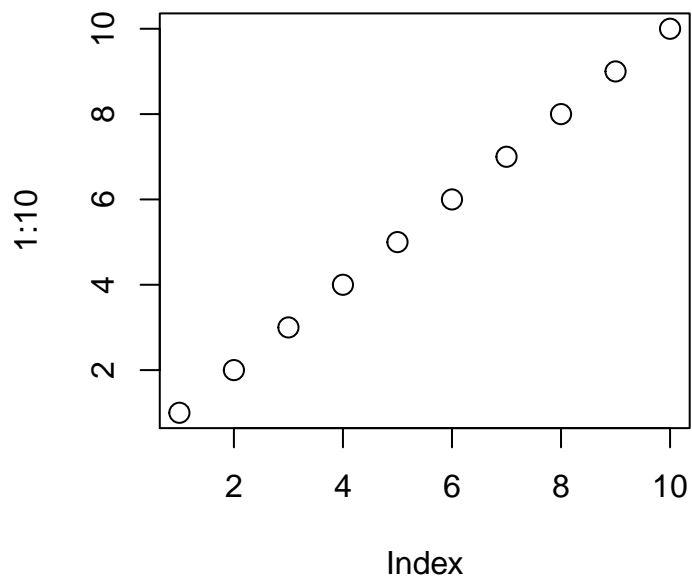
Methods

Results

Problems with binning



correlation of various goodness of fit metrics



relative AIC with sampling

hypothesis tests with sampling

probability of internal modes with sampling and generating model

model distinguishability across NB parameter space

Discussion

- Don't bin
- Don't "average"
- Fit with likelihood
- Sampling is important and the veil line is too simplistic

References