We found evidence that the shapes of the SADs for a range of real ecological communities are statistically unlikely when compared to their feasible sets. For four of the five datasets we analyzed – BBS, Gentry, Mammal Communities, and Misc. Abund – empirical SADs are highly skewed and highly uneven relative to their feasible sets much more frequently than would be expected by chance ([Figure](#_Figure_3:_Skewness_1), [Figure](#_Figure_4:_Evenness)). Combined across these four datasets, 16% of observed SADs are more skewed than 95% of their feasible sets, and 31% are less even than 95% of their feasible sets. By chance we would expect only 5% of observed distributions to fall in these extremes. However, we detected considerably less pronounced deviations for communities from the FIA communities ([Figure](#_Figure_3:_Skewness_1), [Figure](#_Figure_4:_Evenness)), for which percentile scores were near uniformly-distributed for skewness (5% of observations are more skewed than 95% of the feasible set), and much noisier than any of the other datasets for evenness (11.5% of observations are less even). Why the FIA communities show a different pattern is unclear, but may be partially due to their small size, which we explore further below. For the four other datasets, however, our results suggest that nonrandom processes drive observed SADs to be less even than would occur by chance.

These widespread and consistent deviations between empirical SADs and their feasible sets may be the signature of ecological processes operating on top of statistical constraints, and therefore offer new leverage for developing and evaluating theories to predict the shape of the SAD. One logical next step is to test whether existing theories (e.g. neutral theory, METE) and common functional approximations (logseries, exponential, and log normal) can accurately predict deviations from the feasible set. Many of these theoretical predictions generate hollow curves, but have yet to be evaluated on their ability to accurately capture nuanced deviations of the type we document here. In addition, our results suggest that the prevailing processes that structure communities tend to be ones that push abundance distributions towards an uneven state, rather than those that cause individuals to be spread evenly across species. These could be processes that promote the persistence of rare species at extremely low abundances – thereby lengthening the rare tail of the SAD – or processes that encourage dominant species to be hyper-dominant without driving other species entirely to extinction. Identifying the processes at play in particular systems will require further exploration, but the approach we have demonstrated here reveals the signal to try and explain.

We found considerably weaker evidence for deviations for one of our five datasets, perhaps because of statistical issues related to community size. Unlike the other four datasets, communities in the FIA dataset showed at most weak evidence of deviations from their feasible sets. These communities are by far the smallest in our overall database: [most of them] have fewer than [X species/N individuals]; only [x%] of communities from the other four datasets are within this size range. Community size – in terms of S and N – may affect our ability to detect deviations via its potential effect on how similar the elements of the feasible set are to each other. Small communities may not have enough possible arrangements if their subcomponents, or elements in their feasible sets, to generate highly resolved distributions for the most probable shapes. If this occurs, observations may deviate from the most-likely form, but only the most extreme deviations will be highly unlikely given the breadth of the corresponding probability distribution. When we compared the distributions of shape metrics for small communities to those for large ones, we found that samples from the feasible sets for small communities generate broader distributions of evenness, and especially skewness, than those for large communities ([Figure](#_Figure_8:_95%)). For communities of the sizes represented in the FIA dataset, the 95% interval of skewness values often encompasses more than 80% of the entire range of values; for larger communities, the 95% interval spans closer to 60% of the full range. These broad distributions may not be specific enough to constitute a strong statistical expectation, and the deviations – or lack thereof – that we detect are correspondingly less informative. Additionally, we found a similar lack of detectable deviations among the (few) communities from other datasets that fell within the FIA size range (Figure): for small communities not from FIA X% of skewness and X% of evenness were unlikely relative to their feasible sets. Although this comparison is based on relatively few communities (X non-FIA, small, communities), it is consistent with the possibility that small community size, and not other characteristics related to FIA, may drive this phenomenon.

If this is indeed the case, it means that small-community considerations may affect our capacity to meaningfully distinguish signal from randomness for communities with ranges of S and N that are quite common in ecology. As a group, the FIA communities had especially broad distributions of shape metrics, and exhibited an overall lack of detectable signal. These communities range in size from x to y species and x to y individuals. While these are by no means hard thresholds, they may indicate a general range of values below which we have relatively diminished power to distinguish deviation from randomness. Unless we can develop more sensitive methods for identifying deviations even for these small communities, we may stand to learn the most by focusing on SADs from relatively large communities. In the meantime, sampling the range of forms represented in the feasible set helps us identify when the distribution of shapes present is relatively broad, and may even help us develop a type of power analysis to find the smallest-detectable deviation for a given community size.

It is also important to recognize that there are multiple plausible approaches to the defining the statistical baseline for the SAD, of which we have taken only one. Other formulations for the statistical baseline may be equally valid and can generate different statistical expectations, including forms that approximate exponential, Poisson, or log-series distributions. There is currently no unambiguous logical argument for one baseline over another, but comparing the performance of different baselines is clearly an important next step in this process towards reinvigorating the use of the SAD as a diagnostic tool.

Characterizing and adjusting for a statistical baseline for the SAD, as we demonstrate here, refreshes our perspective on the distribution and opens up several new avenues for better understanding how and when biological drivers affect its shape. Persistent deviations between observed communities and their baselines may be evidence of biological processes operating on top of fundamental statistical constraints, and focusing on these deviations could offer new leverage for evaluating theoretical predictions for the SAD. In doing so, we must appreciate that there is considerable nuance to defining the appropriate statistical baseline and calibrating our expected power to detect deviations, especially for small communities. Exploring other constructions for the baseline, and developing methods for establishing if not improving the constraints on our ability to detect deviations, would further clarify how statistical constraints manifest in the SAD and what power we have to disentangle biological signal from randomness. Our results here suggest that statistical constraints have strong effects on the SAD, but that these constraints alone do not fully account for the extremely uneven SADs we observe in nature – leaving an important role for ecological process. Continuing to explore and account for the interplay between statistical constraint and biological process constitutes a promising and profound new approach to our understanding of this familiar, yet surprisingly mysterious, ecological pattern.