

Sampling Feasible Sets

Contents

1 Species Abundance Distribution (SAD) Definition	1
2 Feasible Sets	1
2.1 Alternative Representation	2
2.2 Generating SADs	2
2.3 Counting Feasible Sets	2
2.4 Sampling Feasible Sets	3
References	3

1 Species Abundance Distribution (SAD) Definition

For $n \geq s \geq 1$, we define a SAD as:

$$\mathbf{x} = (x_1, x_2, \dots, x_s)$$

where

$$\begin{aligned} x_i &\in \mathbb{Z}^+ \\ x_i &\leq x_j \text{ for } i \leq j \\ \sum x_i &= n \end{aligned}$$

In effect, a SAD is a division of n individuals among s classes (i.e. taxonomic categories), and we are interested only in the sizes of those classes, disregarding their specific identity (i.e. the classes are interchangeable).

Note that we deviate slightly from ecological convention in that we order the sizes of the classes in ascending, as opposed to descending order.

2 Feasible Sets

A **Feasible Set**, $F(s, n)$, is the set of all possible SADs for a particular s and n , and $f(s, n) = |F(s, n)|$ is the cardinality of $F(s, n)$ (i.e. the size of the Feasible Set). The values of $f(s, n)$ are identical to the number of partitions of n into exactly s parts¹.

¹ $f(s, n)$, the number of partitions of n into exactly s parts is more commonly expressed via the related quantity, $p(s, n)$, the number of partitions of s into at most n parts; the two are related by $f(s, n) = p(s - n, n)$. Some of the earliest calculations of $p(s, n)$ were performed by Euler in a 10 November 1742 letter to Bernoulli, in which values were computed via recurrence relation (1862); for a more extensive history of the topic, see Dickson (1919).

2.1 Alternative Representation

First, we describe an alternative representation for SADs, where we describe the relative differences in sizes of consecutive classes instead of the sizes of the classes directly:

$$\mathbf{y} = (y_1, y_2, \dots, y_s)$$

where

$$\begin{aligned} y_1 &= x_1 \\ y_i &= x_i - x_{i-1} \geq 0 \text{ for } i \geq 2 \end{aligned}$$

Thus,

$$\begin{aligned} x_1 &= y_1 \\ x_2 &= y_1 + y_2 \\ x_3 &= y_1 + y_2 + y_3 \\ &\vdots \\ x_i &= \sum_{j=1}^{j=i} y_j \end{aligned}$$

and

$$\begin{aligned} n &= x_1 + x_2 + \dots + x_s \\ &= s \cdot y_1 + (s-1) \cdot y_2 + \dots + 1 \cdot y_s \end{aligned}$$

2.2 Generating SADs

We describe a generative approach for creating a SAD.

1. Set $i = 1$, the index of the class whose size will be determined next. Set $n_r = n$, the number of individuals remaining to be allocated. Set $s_r = s$, the number of classes remaining to be filled.
2. Choose a value for y_1 . The possible values are $[1, \lfloor \frac{n_r}{s_r} \rfloor]$.
3. Compute n_r , the remaining un-allocated individuals, as $n_r = n_r - s_r \cdot y_i$. Compute s_r , the remaining un-allocated classes, as $s_r = s_r - 1$. Increment $i = i + 1$.
4. Choose a value for y_i . The possible values are $[0, \lfloor \frac{n_r}{s_r} \rfloor]$.
5. Repeat steps 3 and 4, until the SAD is fully generated.

2.3 Counting Feasible Sets

Mirroring the generative approach, we can define $f(s, n)$ recursively.

1. Define $r(s, n)$ to be the number of ways of allocating n individuals among s classes, and allowing the sizes of the classes to include 0.
2. $f(s, n) = \sum_{y_1=1}^{y_1=\lfloor \frac{n}{s} \rfloor} r(s-1, n-s \cdot y_1)$
3. $r(s, n) = \sum_{y_i=0}^{y_i=\lfloor \frac{n}{s} \rfloor} r(s-1, n-s \cdot y_i)$

where $r(1, n) = 1$ (there is only way to allocate n individuals among 1 class).

2.4 Sampling Feasible Sets

We can also sample $F(s, n)$ uniformly, using a similarly recursive approach.

1. For all $y_1 \in [1, \lfloor \frac{n}{s} \rfloor]$, compute $r(s-1, n-s \cdot y_1)$.
2. The sum of these values of r determine the relative probabilities for the corresponding value of y_1 .
3. Choose y_1 based on the corresponding weights, $\frac{r(s-1, n-s \cdot y_1)}{\sum_{y_1} r(s-1, n-s \cdot y_1)}$
4. Choose the remaining y_i according to the corresponding values of $r(s, n)$ as used to count the feasible set.

References

- Dickson, Leonard E. 1919. *History of the Theory of Numbers*. Vol. 2: Diophantine Analysis. Carnegie Institution of Washington. <https://archive.org/details/historyoftheoryo02dickuoft/page/100/mode/2up>.
- Euler, Leonhard. 1742. *Sex Litterae Ad Nicolaum Bernoullium Ii, Basileensem J. U. D. Datae 1742 Ad 1745*. 820. Euler Archive - All Works. <https://scholarlycommons.pacific.edu/euler-works/820>.