Module "4": Lab 1

Ellen Bledsoe

2024-12-03

Species Richness and Diversity

Today, we will be working with community data to calculate species richness, plot species accumulation curves, and calculate two different of species diversity indices.

Set-Up

For today's lab and assignment, we will need to load a new package called `vegan`. I've already installed the package on Posit Cloud, so we only need to load it into our work space.

```r
library(vegan)
```

The `vegan` package is one of the most popular packages available for working with community-level data. It includes functions for many multivariate analyses commonly used in the analysis of ecological communities. We will be using it today for its species accumulation curve functions.

Species Accumulation Curves

For this example, we are going to use a dataset of fish abundances from the Doubs River in France. Each column represents a species, and each row is a survey.

```r
fish <- read.csv("../data_raw/Doubs_fish.csv")
head(fish)
```

```
##   CHA TRU VAI LOC OMB BLA HOT TOX VAN CHE BAR SPI GOU BRO PER BOU PSO ROT CAR
## 1   0   3   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0
## 2   0   5   4   3   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0
## 3   0   5   5   5   0   0   0   0   0   0   0   0   0   1   0   0   0   0   0
## 4   0   4   5   5   0   0   0   0   0   1   0   0   1   2   2   0   0   0   0
## 5   0   2   3   2   0   0   0   0   5   2   0   0   2   4   4   0   0   2   0
## 6   0   3   4   5   0   0   0   0   1   2   0   0   1   1   1   0   0   0   0
##   TAN BCO PCH GRE GAR BBO ABL ANG
## 1   0   0   0   0   0   0   0   0
## 2   0   0   0   0   0   0   0   0
## 3   0   0   0   0   0   0   0   0
## 4   1   0   0   0   0   0   0   0
## 5   3   0   0   0   5   0   0   0
## 6   2   0   0   0   1   0   0   0
```

First, we need to run our survey by species matrix through the `specaccum()` function to prepare the data to be put into a species accumulation curve model.

```
fish_sac <- specaccum(fish)
```

Now, we want to put the object we just created, `fish_sac`, into the `fitspecaccum()` function. We need to specify the model we want as "asymp".

```
fish_fit_sac <- fitspecaccum(fish_sac, model = "asymp")
```

Now that we've fit the model, we can get out the estimated asymptote of the species accumulation curve and plot the predicted curve.

`coef()` gives us the estimated asymptote of the species accumulation curve. The only value you need to focus on is the "Asym" (asymptote) value.

```
coef(fish_fit_sac)
```

```
##        Asym         R0         lrc
## 26.9858415  2.2980754 -0.6184861
```

The asymptote is coming out right around 27, meaning that there are likely about 27 species in our community. If we look at the original data, we see that the data frame has 27 columns (AKA species), so we've probably caught all the species in the community.

Let's plot the data to see what it looks like.

```
plot(fish_fit_sac)
```
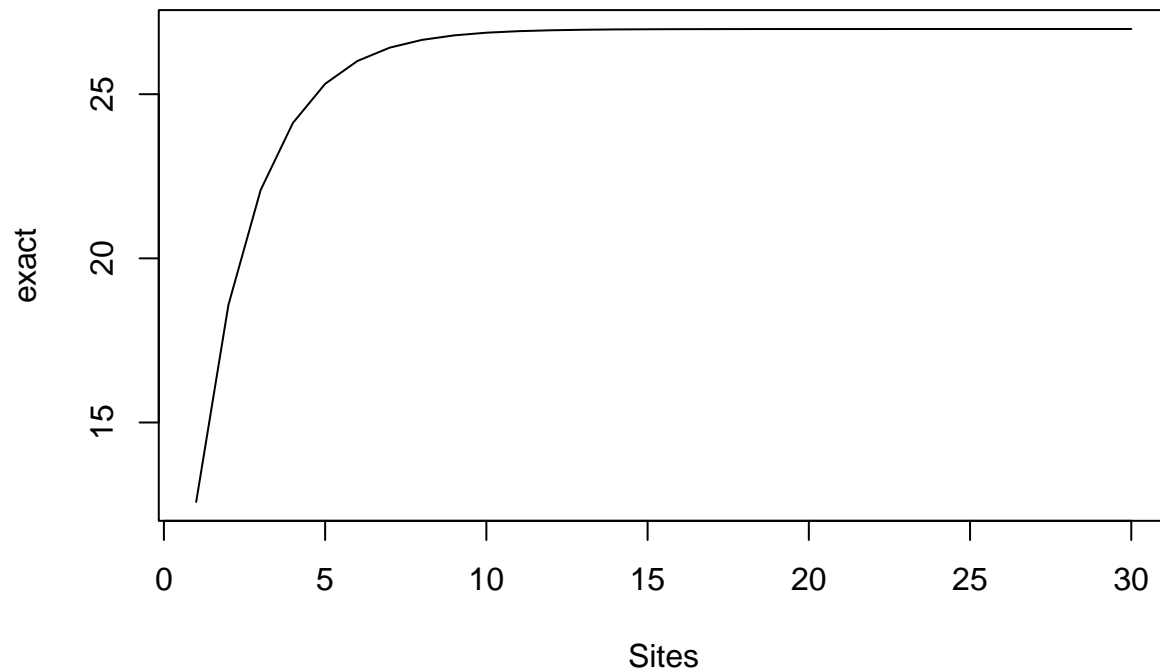

Looks like we hit the asymptote well before we stopped sampling, which is a great sign.

Species Richness and Species Diversity

Let's run through how we can calculate the richness, and Shannon and Simpson Diversity Indices for a community.

```r
# Don't worry about understanding this code!
# I'm making the frequencies of each fish species and making them into a dataframe
# The rownames
fish_abund <- data.frame(abund = fish_sac$freq)
head(fish_abund)
```

```
##     abund
## CHA     8
## TRU    17
## VAI    20
## LOC    24
## OMB     8
## BLA     8
```

Species Richness

Fortunately, calculating species richness is fairly straightforward. We already know that the richness of the fish community is 27, but let's calculate that with code. The `nrow()` function will count how many rows we have for us, which, given the way our data is being presented, will be the species richness of the community.

```r
nrow(fish_abund)
```

```
## [1] 27
```

Shannon Index

We'll start with the Shannon Index for the Doubs fish community.

Our first step is to sum up the total number of individuals in the community. We need this value to create the proportion for each species.

We will use the sum function and specify that we want to add up the numbers in the "value" column by using the dollar sign operator.

```r
total <- sum(fish_abund$abund)
total
```

```
## [1] 375
```

Now we want to calculate the proportion of the population made up by each species. We will store these values by adding them into a new column in the `abund` data frame.

We do this by using the dollar sign operator again. On the left side of the assignment arrow, we create a new column. On the right side, we tell R what the new column should contain. In this case, we want to create the proportion for each species—the number of individuals per species divided by the total number of individuals in the population. We will call this new column `prop`.

```
fish_abund$prop <- fish_abund$abund / total
head(fish_abund)
```

```
##     abund        prop
## CHA     8 0.02133333
## TRU    17 0.04533333
## VAI    20 0.05333333
## LOC    24 0.06400000
## OMB     8 0.02133333
## BLA     8 0.02133333
```

Note that we now a new column with all the proportions calculated. We want to continue this by taking the natural log of that proportion. We use the `log()` function to do that.

```
fish_abund$ln_prop <- log(fish_abund$prop)
head(fish_abund)
```

```
##     abund        prop   ln_prop
## CHA     8 0.02133333 -3.847484
## TRU    17 0.04533333 -3.093713
## VAI    20 0.05333333 -2.931194
## LOC    24 0.06400000 -2.748872
## OMB     8 0.02133333 -3.847484
## BLA     8 0.02133333 -3.847484
```

Now we need to multiple the values in the two columns we just created.

```
fish_abund$prop_x_lnprop <- fish_abund$prop * fish_abund$ln_prop
head(fish_abund)
```

```
##     abund        prop   ln_prop prop_x_lnprop
## CHA     8 0.02133333 -3.847484   -0.08207967
## TRU    17 0.04533333 -3.093713   -0.14024831
## VAI    20 0.05333333 -2.931194   -0.15633033
## LOC    24 0.06400000 -2.748872   -0.17592782
## OMB     8 0.02133333 -3.847484   -0.08207967
## BLA     8 0.02133333 -3.847484   -0.08207967
```

Finally, we can sum these values using the `sum()` function and multiple by -1 to calculate the Shannon Index (H) for this community.

```
H <- sum(fish_abund$prop_x_lnprop) * -1
H
```

```
## [1] 3.239048
```

You will want to use a similar strategy to calculate the Simpson Index, as well.