

Diversity Indices

Pablo E. Gutiérrez-Fonseca

2023-10-01

Load libraries

```
library(vegan)
```

```
## Warning: package 'vegan' was built under R version 4.2.3
```

```
## Loading required package: permute
```

```
## Loading required package: lattice
```

```
## Warning: package 'lattice' was built under R version 4.2.3
```

```
## This is vegan 2.6-4
```

```
library(dplyr)
```

```
## Warning: package 'dplyr' was built under R version 4.2.2
```

```
##
```

```
## Attaching package: 'dplyr'
```

```
## The following objects are masked from 'package:stats':
```

```
##
```

```
## filter, lag
```

```
## The following objects are masked from 'package:base':
```

```
##
```

```
## intersect, setdiff, setequal, union
```

```
library(readxl)
```

```
## Warning: package 'readxl' was built under R version 4.2.2
```

Aquatic Macroinvertebrates in two land uses (Agriculture and Urban streams)

```
data <- read_excel('data.xlsx', sheet = 'diversity')
head(data, 20)
```

```
## # A tibble: 20 x 4
##   site      sample  taxa      abundance
##   <chr>    <chr>   <chr>    <dbl>
## 1 Agriculture sample_3 Annelids      1
## 2 Agriculture sample_3 Tricoptera    1
## 3 Agriculture sample_2 Diptera      2
## 4 Agriculture sample_2 Decapoda      2
## 5 Agriculture sample_1 Tricoptera    4
## 6 Agriculture sample_2 Tricoptera    6
## 7 Agriculture sample_1 Plecoptera   15
## 8 Agriculture sample_3 Plecoptera   19
## 9 Agriculture sample_3 Ephemeroptera 21
## 10 Agriculture sample_1 Ephemeroptera 31
## 11 Agriculture sample_2 Coleoptera   33
## 12 Agriculture sample_2 Ephemeroptera 51
## 13 urban     sample_1 Coleoptera    1
## 14 urban     sample_3 Odonata      1
## 15 urban     sample_3 Decapoda      1
## 16 urban     sample_3 Sphaeriidae    1
## 17 urban     sample_1 Diptera      2
## 18 urban     sample_1 Tricoptera    2
## 19 urban     sample_2 Diptera      2
## 20 urban     sample_2 Megaloptera    2
```

Species richness

Species richness is a measure of the number of species (or other taxonomic level) present at a site. Sites with more taxa are considered **richer**.

```
# Calculate total species richness per site
total_species_richness <- data %>%
  group_by(site) %>%
  summarise(total_species = n_distinct(taxa))

# Display the results
print(total_species_richness)
```

```
## # A tibble: 2 x 2
##   site      total_species
##   <chr>          <int>
## 1 Agriculture      7
## 2 urban            9
```

Abundance

```

# Calculate mean abundance per taxa across all sites
mean_abundance <- data %>%
  group_by(site) %>%
  summarize(mean_abundance = sum(abundance, na.rm = TRUE))

# Display the result
print(mean_abundance)

```

```

## # A tibble: 2 x 2
##   site      mean_abundance
##   <chr>          <dbl>
## 1 Agriculture      186
## 2 urban            65

```

Diversity indices

Shannon_Weaver (or H)

The Shannon entropy (H) is calculated using the formula:

$$H = - \sum_{i=1}^S p_i * \log_b(p_i)$$

where:

- p_i is the proportion of species i ,
- S is the total number of species,
- $\sum_{i=1}^S p_i = 1$ (the sum of all proportions is equal to 1),
- b is the base of the logarithm.

This formula is commonly used to measure the diversity or uncertainty in a system with multiple species, where p_i represents the relative abundance of each species.

- The higher the value of H , the higher the diversity of species in a particular community.
- The lower the value of H , the lower the diversity.
- A value of $H = 0$ indicates a community that only has one species.

Example

```

# Calculate Shannon diversity index
shannon_results <- data %>%
  group_by(site) %>%
  summarise(shannon_index = diversity(c(abundance), index = "shannon"))

# Display the results
print(shannon_results)

```

```

## # A tibble: 2 x 2
##   site      shannon_index
##   <chr>          <dbl>
## 1 Agriculture      1.99
## 2 urban            2.44

```

By hand

```
# Calculat Shannon Diversity
data_agriculture <- data %>%
  filter(site == 'Agriculture')

total_abundance <- sum(data_agriculture$abundance)

proportions_agriculture <- data_agriculture %>%
  mutate(proportion = abundance / total_abundance)

# Calculate Shannon Diversity
shannon_agriculture <- -sum(proportions_agriculture$proportion * log(proportions_agriculture$proportion))
print(paste("Shannon Diversity for Agriculture:", shannon_agriculture))
```

```
## [1] "Shannon Diversity for Agriculture: 1.98956469558767"
```

```
# Calculat Shannon Diversity
data_urban <- data %>%
  filter(site == 'urban')

total_abundance <- sum(data_urban$abundance)

proportions_urban <- data_urban %>%
  mutate(proportion = abundance / total_abundance)

# Calculate Shannon Diversity
shannon_urban <- -sum(proportions_urban$proportion * log(proportions_urban$proportion))
print(paste("Shannon Diversity for Urban:", shannon_urban))
```

```
## [1] "Shannon Diversity for Urban: 2.44231386467766"
```