

Diversity Indices

Pablo E. Gutiérrez-Fonseca

2023-10-01

Load libraries

```
library(vegan)
```

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

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

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

```
library(dplyr)
```

```
##
```

```
## 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)
```

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

```
data <- read_excel('data.xlsx')
```

```
head(data,20)
```

```
## # A tibble: 16 x 3
```

```
##   taxa      abundance site
```

```
##   <chr>          <dbl> <chr>
```

```
## 1 Plecoptera     11.3   Agriculture
```

```
## 2 Diptera        0.667  Agriculture
```

```
## 3 Annelids      0.333 Agriculture
## 4 Ephemeroptera 34.3  Agriculture
## 5 Coleoptera    11    Agriculture
## 6 Tricoptera    3.67  Agriculture
## 7 Decapoda      0.667 Agriculture
## 8 Plecoptera    5.5   Urban
## 9 Diptera       4     Urban
## 10 Ephemeroptera 2     Urban
## 11 Odonata       1     Urban
## 12 Megaloptera   6.33  Urban
## 13 Coleoptera    1     Urban
## 14 Tricoptera    5.33  Urban
## 15 Decapoda      1     Urban
## 16 Sphaeriidae   1     Urban
```

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 species richness per site for species with abundance > 0
species_richness <- data %>%
  filter(abundance > 0) %>%
  group_by(site) %>%
  summarize(species_count = n_distinct(taxa))

# Display the result
print(species_richness)
```

```
## # A tibble: 2 x 2
##   site      species_count
##   <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      62
## 2 Urban           27.2
```

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 entropy per site
shannon_per_site <- data %>%
  group_by(site) %>%
  summarize(shannon_entropy = diversity(abundance, index = "shannon"))

# Display the result
print(shannon_per_site)
```

```
## # A tibble: 2 x 2
##   site      shannon_entropy
##   <chr>          <dbl>
## 1 Agriculture      1.24
## 2 Urban            1.94
```