5. Worksheet: Alpha Diversity

Student Name; Z620: Quantitative Biodiversity, Indiana University

29 January, 2025

OVERVIEW

In this exercise, we will explore aspects of local or site-specific diversity, also known as alpha (α) diversity. First we will quantify two of the fundamental components of (α) diversity: **richness** and **evenness**. From there, we will then discuss ways to integrate richness and evenness, which will include univariate metrics of diversity along with an investigation of the **species abundance distribution (SAD)**.

Directions:

1. In the Markdown version of this document in your cloned repo, change "Student Name" on line 3 (above) to your name.
2. Complete as much of the worksheet as possible during class.
3. Use the handout as a guide; it contains a more complete description of data sets along with the proper scripting needed to carry out the exercise.
4. Answer questions in the worksheet. Space for your answer is provided in this document and indicated by the ">" character. If you need a second paragraph be sure to start the first line with ">". You should notice that the answer is highlighted in green by RStudio (color may vary if you changed the editor theme).
5. Before you leave the classroom, **push** this file to your GitHub repo.
6. For the assignment portion of the worksheet, follow the directions at the bottom of this file.
7. When you are done, **Knit** the text and code into a PDF file.
8. After Knitting, submit the completed exercise by creating a **pull request** via GitHub. Your pull request should include this file `AlphaDiversity_Worskheet.Rmd` and the PDF output of `Knitr` (`AlphaDiversity_Worskheet.pdf`).

1) R SETUP

In the R code chunk below, please provide the code to: 1) Clear your R environment, 2) Print your current working directory, 3) Set your working directory to your `Week-2/` folder folder, and 4) Load the `vegan` R package (be sure to install first if you have not already).

```
rm(list=ls())
getwd()
```

```
## [1] "/cloud/project/QB2025_Park/Week2-Alpha"
```

```
setwd("/cloud/project/QB2025_Park/Week2-Alpha")
```

2) LOADING DATA

In the R code chunk below, do the following: 1) Load the BCI dataset, and 2) Display the structure of the dataset (if the structure is long, use the `max.level = 0` argument to show the basic information).

```
#install.packages("vegan")
require("vegan")
```

```
## Loading required package: vegan
```

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

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

```
## This is vegan 2.6-8
```

```
data(BCI)
str(BCI, max.level=0)
```

```
## 'data.frame':    50 obs. of  225 variables:
##  - attr(*, "original.names")= chr [1:225] "Abarema.macradenium" "Acacia.melanoceras" "Acalypha.divers
```

3) SPECIES RICHNESS

Species richness (S) refers to the number of species in a system or the number of species observed in a sample.

Observed richness

In the R code chunk below, do the following:

1. Write a function called `S.obs` to calculate observed richness
2. Use your function to determine the number of species in `site1` of the BCI data set, and
3. Compare the output of your function to the output of the `specnumber()` function in `vegan`.

```
s.obs <- function(x=""){
rowSums(x>0) * 1
}

s.obs(BCI[1,])
```

```
##  1
## 93
```

```
specnumber(BCI[1,])
```

```
##  1
## 93
```

```
s.obs (BCI[(1:4),])
```

```
##  1  2  3  4
## 93 84 90 94
```

Question 1: Does `specnumber()` from `vegan` return the same value for observed richness in `site1` as our function `S.obs`? What is the species richness of the first four sites (i.e., rows) of the BCI matrix?

> ***Answer 1***: Yes! 1=93, 2=84, 3=90, 4=94

Coverage: How well did you sample your site?

In the R code chunk below, do the following:

1. Write a function to calculate Good's Coverage, and
2. Use that function to calculate coverage for all sites in the BCI matrix.

```
goods<-function(x=""){
  1-(rowSums(x==1)/rowSums(x))
}

goods(BCI)
```

```
##         1         2         3         4         5         6         7         8
## 0.9308036 0.9287356 0.9200864 0.9468504 0.9287129 0.9174757 0.9326923 0.9443155
##         9        10        11        12        13        14        15        16
## 0.9095355 0.9275362 0.9152120 0.9071038 0.9242054 0.9132420 0.9350649 0.9267735
##        17        18        19        20        21        22        23        24
## 0.8950131 0.9193084 0.8891455 0.9114219 0.8946078 0.9066986 0.8705882 0.9030612
##        25        26        27        28        29        30        31        32
## 0.9095023 0.9115479 0.9088729 0.9198966 0.8983516 0.9221053 0.9382423 0.9411765
##        33        34        35        36        37        38        39        40
## 0.9220183 0.9239374 0.9267887 0.9186047 0.9379310 0.9306488 0.9268868 0.9386503
##        41        42        43        44        45        46        47        48
## 0.8880597 0.9299517 0.9140049 0.9168704 0.9234234 0.9348837 0.8847059 0.9228916
##        49        50
## 0.9086651 0.9143519
```

Question 2: Answer the following questions about coverage:

a. What is the range of values that can be generated by Good's Coverage?
b. What would we conclude from Good's Coverage if n_i equaled N?
c. What portion of taxa in `site1` was represented by singletons?
d. Make some observations about coverage at the BCI plots.

Answer 2a: 0-1

Answer 2b: 0

Answer 2c: 0.0691964

Answer 2d: The coverage at all sites is higher than 0.86 so I would conclude that sampling coverage was high at all sites (though I kind of understand this biologically, doesn't depending on singletons to calculate coverage depend on some assumptions about singletons and the environment they were taken from?)

Estimated richness

In the R code chunk below, do the following:

1. Load the microbial dataset (located in the `Week-2/data` folder),
2. Transform and transpose the data as needed (see handout),
3. Create a new vector (`soilbac1`) by indexing the bacterial OTU abundances of any site in the dataset,
4. Calculate the observed richness at that particular site, and
5. Calculate coverage of that site

```
soilbac<-read.table("data/soilbac.txt", sep ="\t", header = TRUE, row.names=1)
soilbac.t<-as.data.frame(t(soilbac))
soilbac1<-soilbac.t[4,]
s.obs(soilbac1)
```

```
## T7_1
## 1416
```

```
goods(soilbac1)
```

```
##      T7_1
## 0.6706855
```

```
#and then just because I'm curious
s.obs(soilbac.t)
```

```
## T1_1 T1_2 T1_3 T7_1 T7_2 T7_3 DF_1 DF_2 CF_1 CF_2 CF_3
## 1074 1302 1174 1416 1406 1143 1806 1151  924 1122  851
```

```
goods(soilbac.t)
```

```
##      T1_1      T1_2      T1_3      T7_1      T7_2      T7_3      DF_1      DF_2
## 0.6479471 0.6676558 0.6735097 0.6706855 0.7342490 0.7812613 0.6406384 0.7779733
##      CF_1      CF_2      CF_3
## 0.8211009 0.6621102 0.7539062
```

Question 3: Answer the following questions about the soil bacterial dataset.

a. How many sequences did we recover from the sample `soilbac1`, i.e. N?
b. What is the observed richness of `soilbac1`?
c. How does coverage compare between the BCI sample (`site1`) and the KBS sample (`soilbac1`)?

Answer 3a: 2903

Answer 3b: 1416

Answer 3c: Although the richness is greater in the KBS sample, coverage was greater in BCI

Richness estimators

In the R code chunk below, do the following:

1. Write a function to calculate **Chao1**,
2. Write a function to calculate **Chao2**,
3. Write a function to calculate **ACE**, and
4. Use these functions to estimate richness at `site1` and `soilbac1`.

```
s.chao1<-function(x=""){
  s.obs(x) + (sum(x==1)^2)/(2*sum(x==2))
}

s.chao2<-function(site="", SbyS=""){
  SbyS=as.data.frame(SbyS)
  x=SbyS[site,]
  SbyS.pa<-(SbyS>0)*1 #converting to presence/absence
  q1=sum(colSums(SbyS.pa)==1) #singletons
  q2=sum(colSums(SbyS.pa)==2) #doubletons
  s.chao2=s.obs(x)+(q1^2)/(2*q2)
  return(s.chao2)
}

s.ace<-function(x="", thresh=10){
  x<-x[x>0] # excludes zero-abundance taxa
  s.abund<-length(which(x>thresh)) # richness of abundant taxa
  s.rare<-length(which(x<=thresh)) # richness of rare taxa
  singlt<-length(which(x==1)) # number of singleton taxa
```

```
  n.rare<-sum(x[which(x<=thresh)]) # abundance of rare individuals
  c.ace<-1-(singlt/n.rare) # coverage (prop non-singlt rare inds)
  i<-c(1:thresh) # threshold abundance range
  count<-function(i,y){ # counter to go through i range
    length(y[y==i])
  }
  a.1<-sapply(i,count,x) # number of individuals in richness i richness classes
  f.1<-(i*(i-1))*a.1 # k(k-1)kf sensu Gotelli
  g.ace<-(s.rare/c.ace)*(sum(f.1)/(n.rare*(n.rare-1)))
  s.ace<-s.abund+(s.rare/c.ace) + (singlt/c.ace) * max(g.ace,0)
  return(s.ace)
}

#site1
site1<-BCI[1,]
s.chao1(site1)
```

```
##        1
## 119.6944
```

```
s.chao2(1, BCI)
```

```
##        1
## 104.6053
```

```
s.ace(site1)
```

```
## [1] 159.3404
```

```
#soilbac1
s.chao1(soilbac1)
```

```
##     T7_1
## 3368.855
```

```
s.chao2(1, soilbac.t)
```

```
##     T1_1
## 21055.39
```

```
s.ace(soilbac1)
```

```
## [1] 5376.43
```

Question 4: What is the difference between ACE and the Chao estimators? Do the estimators give consistent results? Which one would you choose to use and why?

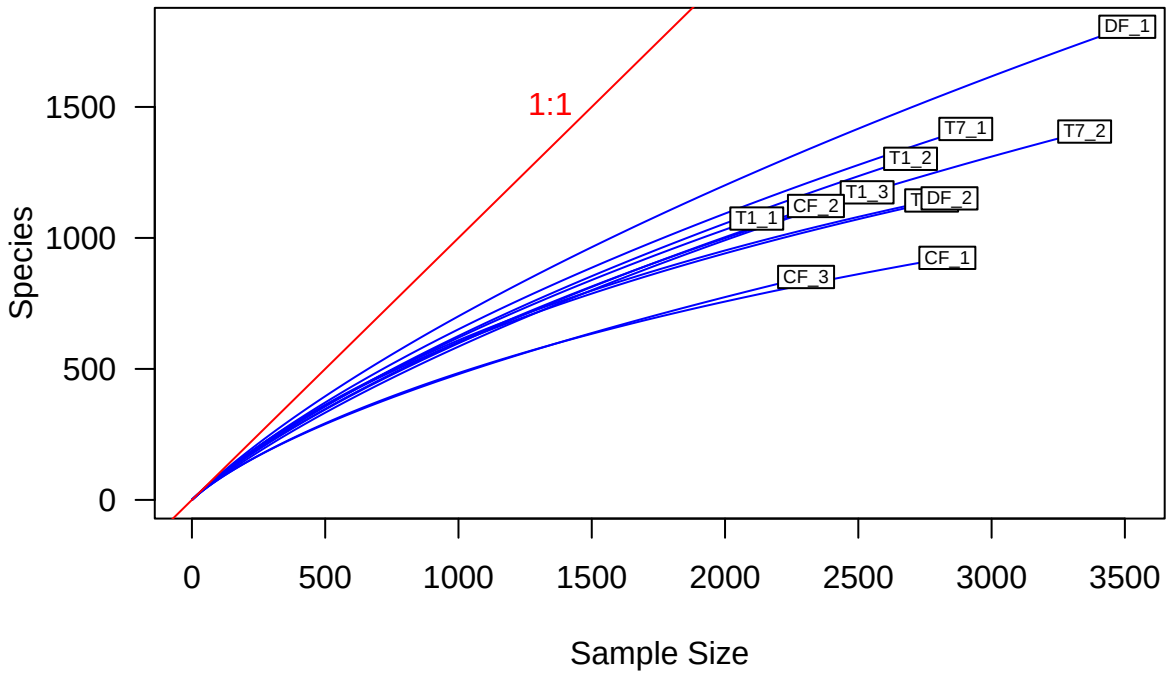
> ***Answer 4***: ACE estimations focus on the abundance of rare species, whereas the chao estimators focus on the presence of rare species. For BCI site 1, Chao1 says we estimate about 120 species, Chao2 gives ~105, and ACE gives ~159. In the soilbac1 vector, we have Chao1 estimating 3369 species, Chao2 estimating 21055 species, and ACE estimating 5376 species. These are not always consistent, though it seems they can be based on the BCI data. It is highly dependent on the presence and abundance of rare species, as shown by the high discrepancy between the Chao2 estimation of richness in the soilbac data. ACE is good for mixed abundance datasets, Chao1 for abundance data with many rare species, and Chao2 for incidence data with many rare species.

Rarefaction

In the R code chunk below, please do the following:

1. Calculate observed richness for all samples in `soilbac`,
2. Determine the size of the smallest sample,
3. Use the `rarefy()` function to rarefy each sample to this level,
4. Plot the rarefaction results, and
5. Add the 1:1 line and label.

```
soilbac.s<-s.obs(soilbac.t)
min.N<-min(rowSums(soilbac.t))
s.rarefy<-rarefy(x=soilbac.t, sample=min.N, se=TRUE)
rarecurve(x=soilbac.t, step=20, col="blue",cex=0.6, las=1)
abline(0, 1, col='red')
text(1500, 1500, "1:1", pos=2, col='red')
```



4) SPECIES EVNENNESS

Here, we consider how abundance varies among species, that is, **species evenness**.

Visualizing evenness: the rank abundance curve (RAC)

One of the most common ways to visualize evenness is in a **rank-abundance curve** (sometime referred to as a rank-abundance distribution or Whittaker plot). An RAC can be constructed by ranking species from the most abundant to the least abundant without respect to species labels (and hence no worries about 'ties' in abundance).

In the R code chunk below, do the following:

1. Write a function to construct a RAC,
2. Be sure your function removes species that have zero abundances,
3. Order the vector (RAC) from greatest (most abundant) to least (least abundant), and
4. Return the ranked vector

```
rac<-function(x=""){
  x.ab=x[x>0]
  x.ab.ranked=x.ab[order(x.ab, decreasing=TRUE)]
  as.data.frame(lapply(x.ab.ranked, unlist))
  return(x.ab.ranked)
  }
```

Now, let us examine the RAC for `site1` of the BCI data set.

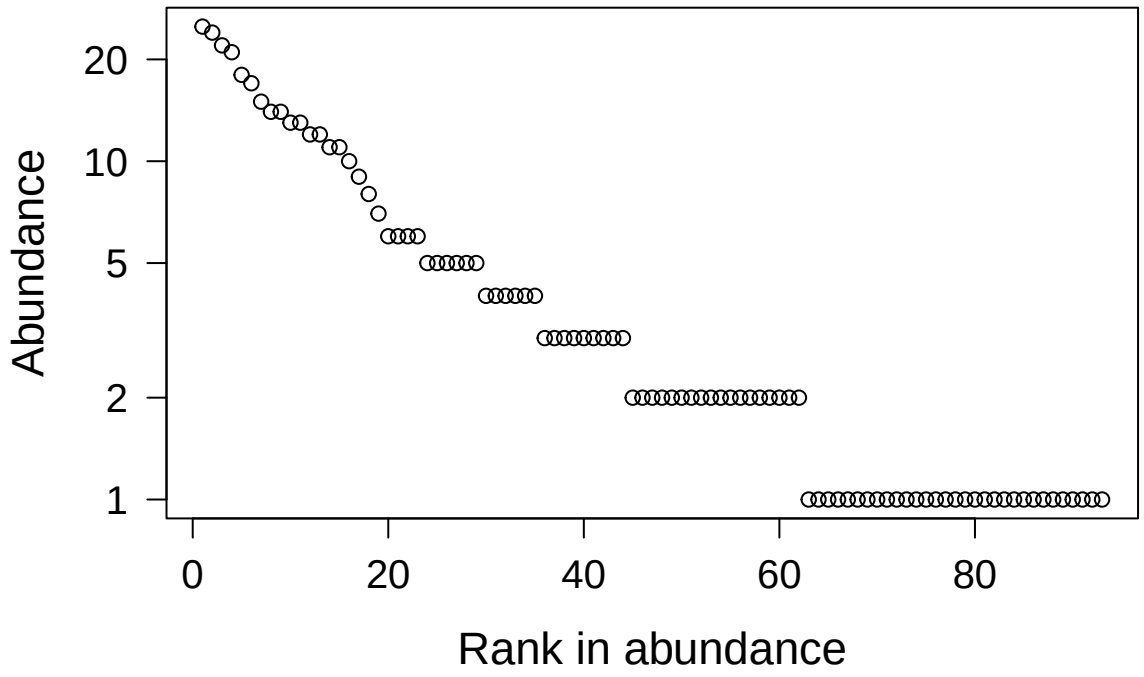
In the R code chunk below, do the following:

1. Create a sequence of ranks and plot the RAC with natural-log-transformed abundances,
2. Label the x-axis "Rank in abundance" and the y-axis "log(abundance)"

```
plot.new()
site1<-BCI[1,]

s1rac<-rac(x=site1)
ranks<-as.vector(seq(1,length(s1rac)))
opar<-par(no.readonly=TRUE)
par(mar=c(5.1, 5.1, 4.1, 2.1))
plot(ranks, log(s1rac), type='p', axes=F,
     xlab="Rank in abundance", ylab="Abundance",
     las=1, cex.lab=1.4, cex.axis=1.25)

box()
axis(side=1, labels=T, cex.axis=1.25)
axis(side=2, las=1, cex.axis=1.25,
     labels=c(1, 2, 5, 10, 20), at=log(c(1, 2, 5, 10, 20)))
```




```
par<-opar
```

Question 5: What effect does visualizing species abundance data on a log-scaled axis have on how we interpret evenness in the RAC?

Answer 5: A log-scaled axis allows us to see species with vastly different abundances on the same plot. We can also see rare species with more clarity, making the tail end of the RAC more pronounced.

Now that we have visualized unevennes, it is time to quantify it using Simpson's evenness ($E_{1/D}$) and Smith and Wilson's evenness index (E_{var}).

Simpson's evenness ($E_{1/D}$)

In the R code chunk below, do the following:

1. Write the function to calculate $E_{1/D}$, and
2. Calculate $E_{1/D}$ for `site1`.

```
simpE<-function(x=""){
  S<-s.obs(x)
  x=as.data.frame(x)
  D<-diversity(x,"inv")
  E<-(D)/S
  return(E)
  }

simpE(site1)
```

```
##         1
## 0.4238232
```

Smith and Wilson's evenness index (E_{var})

In the R code chunk below, please do the following:

1. Write the function to calculate E_{var},
2. Calculate E_{var} for `site1`, and
3. Compare $E_{1/D}$ and E_{var}.

```
Evar<-function(x){
  x<-as.vector(x[x>0])
  1-(2/pi)*atan(var(log(x)))
}

Evar(site1)
```

```
## [1] 0.5067211
```

```
simpE(site1)
```

```
##         1
## 0.4238232
```

Question 6: Compare estimates of evenness for `site1` of BCI using $E_{1/D}$ and E_{var}. Do they agree? If so, why? If not, why? What can you infer from the results?

Answer 6: Simpson's evenness gives us a value of 0.424 and E_{var} gives us 0.507. They seem similar to me (though I'm not familiar with this type of measurement, which makes me wonder if a difference of 0.1 is a lot or not at all), but E_{var} indicates slightly higher evenness. ($E_{1/D}$) is biased by dominant species (by measuring probability of two individuals belonging to the same taxa) while E_{var} is less so because it considers variance of abundance. In consideration of this, it seems that there are a few very abundant species causing ($E_{1/D}$) to be lower than E_{var}.

5) INTEGRATING RICHNESS AND EVENNESS: DIVERSITY METRICS

So far, we have introduced two primary aspects of diversity, i.e., richness and evenness. Here, we will use popular indices to estimate diversity, which explicitly incorporate richness and evenness. We will write our own diversity functions and compare them against the functions in `vegan`.

Shannon's diversity (a.k.a., Shannon's entropy)

In the R code chunk below, please do the following:

1. Provide the code for calculating H' (Shannon's diversity),
2. Compare this estimate with the output of `vegan`'s diversity function using method = "shannon".

```
shanH<-function(x =""){
  H=0
  for(n_i in x){
    if(n_i > 0){
      p=n_i/sum(x)
      H=H-p*log(p)
    }
  }
  return(H)
}

shanH(site1)
```

```
## [1] 4.018412
```

```
diversity(site1, index="shannon")
```

```
## [1] 4.018412
```

Simpson's diversity (or dominance)

In the R code chunk below, please do the following:

1. Provide the code for calculating D (Simpson's diversity),
2. Calculate both the inverse (1/D) and 1 - D,
3. Compare this estimate with the output of `vegan's` diversity function using method = "simp".

```
simpD<-function(x=""){
  D=0
  N=sum(x)
  for(n_i in x){
    D=D+(n_i^2)/(N^2)
  }
  return(D)
}
D.inv <- 1/simpD(site1)
D.sub <- 1-simpD(site1)
print(D.inv)
```

```
## [1] 39.41555
```

```
print(D.sub)
```

```
## [1] 0.9746293
```

```
diversity(site1, "inv")
```

```
## [1] 39.41555
```

```
diversity(site1, "simp")
```

```
## [1] 0.9746293
```

Fisher's α

In the R code chunk below, please do the following:

1. Provide the code for calculating Fisher's α,
2. Calculate Fisher's α for `site1` of BCI.

```
racv<-as.vector(site1[site1>0])
fisher<-fisher.alpha(racv)
fisher
```

```
## [1] 35.67297
```

Question 7: How is Fisher's α different from $E_{H'}$ and E_{var}? What does Fisher's α take into account that $E_{H'}$ and E_{var} do not?

> ***Answer 7***: Fisher's α primarily focuses on richness as opposed to $E_{H'}$ and E_{var}, which focus more on evenness (though Shannon's diversity also considers richness). Fisher's α also specifically assumes a log series distribution, whereas Shannon's diversity and Fisher's α are more general with no assumptions of distribution.

6) HILL NUMBERS

Remember that we have learned about the advantages of Hill Numbers to measure and compare diversity among samples. We also learned to explore the effects of rare species in a community by examining diversity for a series of exponents q.

Question 8: Using `site1` of BCI and `vegan` package, a) calculate Hill numbers for q exponent 0, 1 and 2 (richness, exponential Shannon's entropy, and inverse Simpson's diversity). b) Interpret the effect of rare species in your community based on the response of diversity to increasing exponent q.

Answer 8a:

```
q0<-s.obs(site1)
q1<-exp(diversity(site1, index="shannon"))
q2<-diversity(site1, index="invsimpson")
print(q0)
```

```
##  1
## 93
```

```
print(q1)
```

```
## [1] 55.6127
```

```
print(q2)
```

```
## [1] 39.41555
```

> ***Answer 8b***: Q0 gives equal weight to all species, q1 measures evenness with some sensitivity to rare species, and q2 gives more weight to common species over rare species. Because q2 for site1 is lower than q1, we can infer that site 1 is dominated by few abundant species.

##7) MOVING BEYOND UNIVARIATE METRICS OF α DIVERSITY

The diversity metrics that we just learned about attempt to integrate richness and evenness into a single, univariate metric. Although useful, information is invariably lost in this process. If we go back to the rank-abundance curve, we can retrieve additional information – and in some cases – make inferences about the processes influencing the structure of an ecological system.

Species abundance models

The RAC is a simple data structure that is both a vector of abundances. It is also a row in the site-by-species matrix (minus the zeros, i.e., absences).

Predicting the form of the RAC is the first test that any biodiversity theory must pass and there are no less than 20 models that have attempted to explain the uneven form of the RAC across ecological systems.

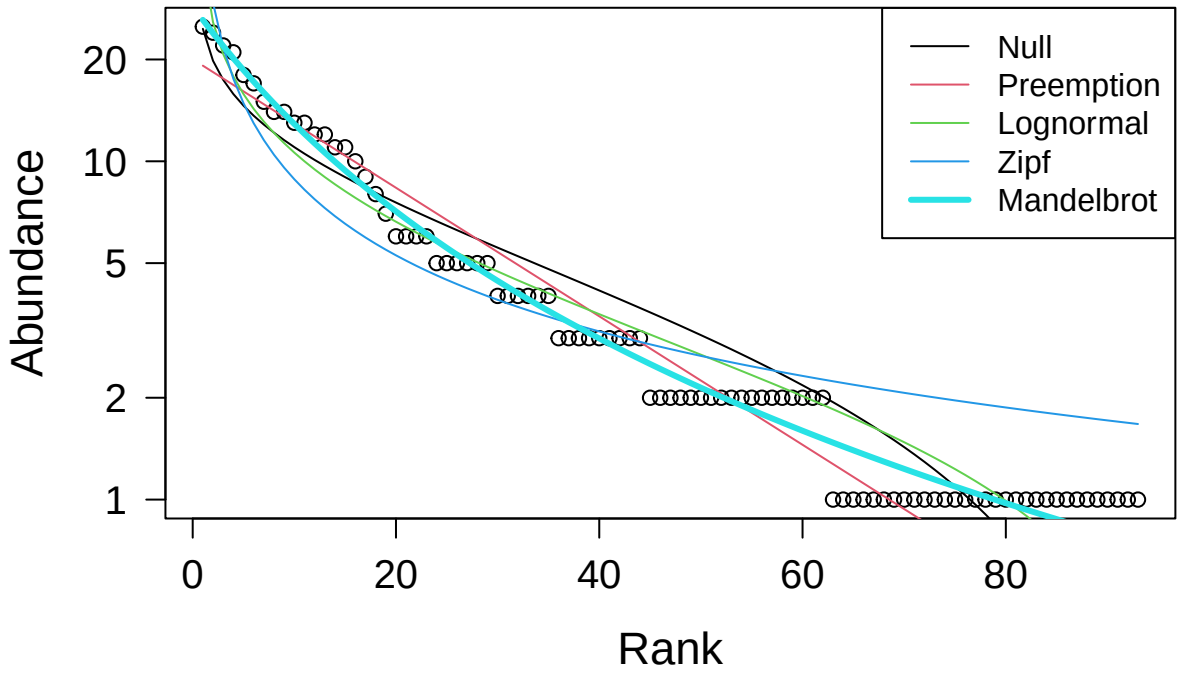
In the R code chunk below, please do the following:

1. Use the `radfit()` function in the `vegan` package to fit the predictions of various species abundance models to the RAC of `site1` in BCI,
2. Display the results of the `radfit()` function, and
3. Plot the results of the `radfit()` function using the code provided in the handout.

```
RACresults<-radfit(site1)
print(RACresults)
```

```
##
## RAD models, family poisson
## No. of species 93, total abundance 448
##
##            par1      par2     par3    Deviance AIC      BIC
## Null                                    39.5261 315.4362 315.4362
## Preemption  0.042797                     21.8939 299.8041 302.3367
## Lognormal   1.0687    1.0186            25.1528 305.0629 310.1281
## Zipf        0.11033  -0.74705           61.0465 340.9567 346.0219
## Mandelbrot  100.52   -2.312    24.084    4.2271 286.1372 293.7350
```

```
plot.new()
plot(RACresults, las=1, cex.lab=1.4, cex.axis=1.25)
```



Question 9: Answer the following questions about the rank abundance curves: a) Based on the output of `radfit()` and plotting above, discuss which model best fits our rank-abundance curve for `site1`? b) Can we make any inferences about the forces, processes, and/or mechanisms influencing the structure of our system, e.g., an ecological community?

> ***Answer 9a***: The Mandelbrot model seems to best fit the RAC for site 1, confirmed by the output of radfit(). The AIC and BIC values are lowest in the Mandelbrot model, indicating best fit.
>
> ***Answer 9b***: Since it seems that site1 is dominated by a few highly abundant species, we could make an inference that the location and/or community at this site is influenced in a way that few species are able to dominate. For example, maybe the location of the community is physically shaped in a way that is more beneficial to plants that have shallow roots, or the nutrients in the soil are more suited for certain species.

Question 10: Answer the following questions about the preemption model: a. What does the preemption model assume about the relationship between total abundance (N) and total resources that can be preempted? b. Why does the niche preemption model look like a straight line in the RAD plot?

> ***Answer 10a***: This model assumes a sequential colonization where each species takes the same fixed proportion of the remaining resources, limiting abundance as more species arrive. ***Answer 10b***: It looks like a straight line because there is an exponential decrease in abundance, leading to a straight line in the log scale.

Question 10: Why is it important to account for the number of parameters a model uses when judging how well it explains a given set of data?

> ***Answer 11***: When models have too many parameters, it can lead to overfitting of training data that doesn't fit new data or overcomplication that makes data difficult to interpret. When they have too few, it may not capture underlying patterns but may be easier to interpret.

SYNTHESIS

1. As stated by Magurran (2004) the $D = \sum p_i^2$ derivation of Simpson's Diversity only applies to communities of infinite size. For anything but an infinitely large community, Simpson's Diversity index is calculated as $D = \sum \frac{n_i(n_i-1)}{N(N-1)}$. Assuming a finite community, calculate Simpson's D, 1 - D, and

Simpson's inverse (i.e. 1/D) for `site 1` of the BCI site-by-species matrix.

```
simpD_finite<-function(x=""){
  D=0
  N=sum(x)
  for(n_i in x){
    D=D+(n_i*(n_i-1))/(N*(N-1))
  }
  return(D)
}

D.inv_finite <- 1/simpD_finite(site1)
D.sub_finite <- 1-simpD_finite(site1)
print(D.inv)
```

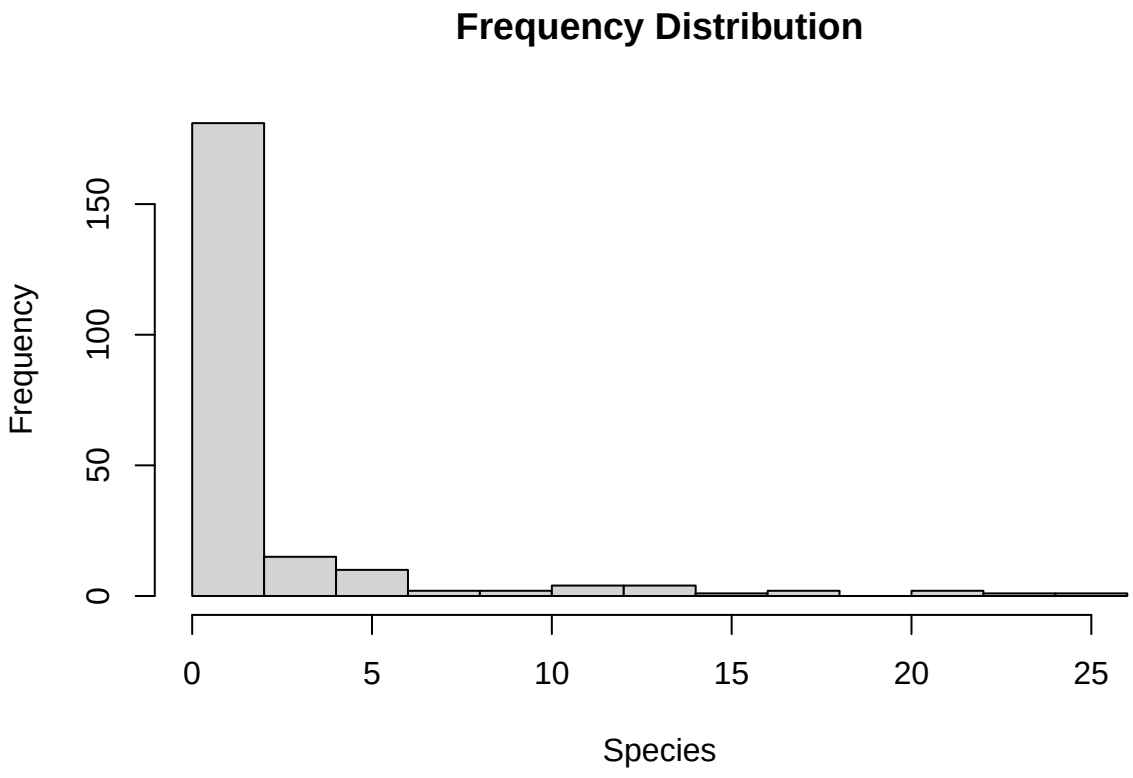
```
## [1] 39.41555
```

```
print(D.sub)
```

```
## [1] 0.9746293
```

2. Along with the rank-abundance curve (RAC), another way to visualize the distribution of abundance among species is with a histogram (a.k.a., frequency distribution) that shows the frequency of different abundance classes. For example, in a given sample, there may be 10 species represented by a single individual, 8 species with two individuals, 4 species with three individuals, and so on. In fact, the rank-abundance curve and the frequency distribution are the two most common ways to visualize the species-abundance distribution (SAD) and to test species abundance models and biodiversity theories. To address this homework question, use the R function **hist()** to plot the frequency distribution for `site 1` of the BCI site-by-species matrix, and describe the general pattern you see.

```
site1 <- as.numeric(site1)
hist(site1, main="Frequency Distribution", xlab="Species", ylab="Frequency")
```



3. We asked you to find a biodiversity dataset with your partner. This data could be one of your own or it could be something that you obtained from the literature. Load that dataset.

```
zoobs<-read.csv("data/sps_macrozoobenthos_allyear.csv")
zoobs <- as.data.frame(zoobs)
zoobs1 <- zoobs[,2:50]
s.obs(zoobs1)
```

```
##  [1] 11  5  7 32 20 17  1 29 10  8  4 27  7 13 13  5  2  0  1
```

```
goods(zoobs1)
```

```
##  [1] 1.0000000 1.0000000 1.0000000 0.9997751 0.9964897 0.9999133 1.0000000
##  [8] 0.9998014 0.9990632 1.0000000 1.0000000 0.9996833 1.0000000 0.9995893
## [15] 0.0000000 1.0000000 1.0000000       NaN 1.0000000
```

```
s.chao1(zoobs1)
```

```
##  [1] 137.75 131.75 133.75 158.75 146.75 143.75 127.75 155.75 136.75 134.75
## [11] 130.75 153.75 133.75 139.75 139.75 131.75 128.75 126.75 127.75
```

```
s.chao2(, zoobs1)
```

```
## NA
## NA
```

```
s.ace(zoobs1)
```

```
## [1] 293.1155
```

How many sites are there? >19 sites

How many species are there in the entire site-by-species matrix? >50 species

Any other interesting observations based on what you learned this week? >Strangely enough, I don't think I have any doubletons since chao2 isn't working (and based off the glance of the dataset). I also find it interesting that my estimated richness is much higher than my observed richness despite coverage being great (via Goods), however this may be due to the fact that the SbyS matrix is a conglomeration of all the individuals across 25 years.

SUBMITTING YOUR ASSIGNMENT

Use Knitr to create a PDF of your completed 5.AlphaDiversity_Worksheet.Rmd document, push it to GitHub, and create a pull request. Please make sure your updated repo include both the pdf and RMarkdown files.

Unless otherwise noted, this assignment is due on **Wednesday, January 29th, 2025 at 12:00 PM (noon)**.