5. Worksheet: Alpha Diversity

Yongsoo Choi; 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_Choi/Week2-Alpha"
```

```
setwd("/cloud/project/QB2025_Choi/Week2-Alpha/")
install.packages("vegan")
```

```
## Installing package into '/cloud/lib/x86_64-pc-linux-gnu-library/4.4'
## (as 'lib' is unspecified)
```

```
require("vegan")
```

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

```
## Loading required package: permute
## Loading required package: lattice
## This is vegan 2.6-8
```

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).

```
data(BCI)
head(BCI)
```

```
##   Abarema.macradenia Vachellia.melanoceras Acalypha.diversifolia
## 1                  0                     0                     0
## 2                  0                     0                     0
## 3                  0                     0                     0
## 4                  0                     0                     0
## 5                  0                     0                     0
## 6                  0                     0                     0
##   Acalypha.macrostachya Adelia.triloba Aegiphila.panamensis
## 1                     0              0                    0
## 2                     0              0                    0
## 3                     0              0                    0
## 4                     0              3                    0
## 5                     0              1                    1
## 6                     0              0                    0
##   Alchornea.costaricensis Alchornea.latifolia Alibertia.edulis
## 1                       2                   0               0
## 2                       1                   0               0
## 3                       2                   0               0
## 4                      18                   0               0
## 5                       3                   0               0
## 6                       2                   1               0
##   Allophylus.psilospermus Alseis.blackiana Amaioua.corymbosa
## 1                       0               25                 0
## 2                       0               26                 0
## 3                       0               18                 0
## 4                       0               23                 0
## 5                       1               16                 0
## 6                       0               14                 0
##   Anacardium.excelsum Andira.inermis Annona.spraguei Apeiba.glabra
## 1                   0              0               1            13
## 2                   0              0               0            12
## 3                   0              0               1             6
## 4                   0              0               0             3
## 5                   0              1               0             4
## 6                   0              1               0            10
##   Apeiba.tibourbou Aspidosperma.desmanthum Astrocaryum.standleyanum
## 1                2                       0                       0
## 2                0                       0                       2
## 3                1                       0                       1
## 4                1                       1                       5
## 5                0                       1                       6
## 6                0                       1                       2
```

```
##   Astronium.graveolens Attalea.butyracea Banara.guianensis
## 1                    6                 0                 0
## 2                    0                 1                 0
## 3                    1                 0                 0
## 4                    3                 0                 0
## 5                    0                 0                 0
## 6                    1                 1                 0
##   Beilschmiedia.pendula Brosimum.alicastrum Brosimum.guianense
## 1                     4                   5                  0
## 2                     5                   2                  0
## 3                     7                   4                  0
## 4                     5                   3                  0
## 5                     8                   2                  0
## 6                     6                   2                  0
##   Calophyllum.longifolium Casearia.aculeata Casearia.arborea
## 1                       0                 0                1
## 2                       2                 0                1
## 3                       0                 0                3
## 4                       2                 0                2
## 5                       1                 0                4
## 6                       2                 0                1
##   Casearia.commersoniana Casearia.guianensis Casearia.sylvestris
## 1                      0                   0                   2
## 2                      0                   0                   1
## 3                      1                   0                   0
## 4                      0                   0                   0
## 5                      1                   0                   0
## 6                      0                   0                   3
##   Cassipourea.guianensis Cavanillesia.platanifolia Cecropia.insignis
## 1                      2                         0               12
## 2                      0                         0                5
## 3                      1                         0                7
## 4                      1                         0               17
## 5                      3                         0               21
## 6                      4                         0                4
##   Cecropia.obtusifolia Cedrela.odorata Ceiba.pentandra Celtis.schippii
## 1                    0               0               0              0
## 2                    0               0               1              0
## 3                    0               0               1              0
## 4                    0               0               0              2
## 5                    1               0               1              2
## 6                    0               0               0              0
##   Cespedesia.spathulata Chamguava.schippii Chimarrhis.parviflora
## 1                     0                  0                    0
## 2                     0                  0                    0
## 3                     0                  0                    0
## 4                     0                  0                    0
## 5                     0                  0                    0
## 6                     0                  0                    0
##   Maclura.tinctoria Chrysochlamys.eclipes Chrysophyllum.argenteum
## 1                 0                     0                      4
## 2                 0                     0                      1
## 3                 0                     0                      2
## 4                 0                     0                      2
```

```
## 5                 0                    0                    6
## 6                 0                    0                    2
##   Chrysophyllum.cainito Coccoloba.coronata Coccoloba.manzinellensis
## 1                     0                  0                       0
## 2                     0                  0                       0
## 3                     0                  0                       0
## 4                     0                  1                       0
## 5                     0                  2                       0
## 6                     0                  0                       0
##   Colubrina.glandulosa Cordia.alliodora Cordia.bicolor Cordia.lasiocalyx
## 1                    0                2             12                 8
## 2                    0                3             14                 6
## 3                    0                3             35                 6
## 4                    0                7             23                11
## 5                    0                1             13                 7
## 6                    0                1              7                 6
##   Coussarea.curvigemma Croton.billbergianus Cupania.cinerea Cupania.latifolia
## 1                    0                    2               0                 0
## 2                    0                    2               0                 0
## 3                    0                    0               0                 0
## 4                    1                   11               0                 1
## 5                    0                    6               0                 0
## 6                    2                    0               0                 0
##   Cupania.rufescens Cupania.seemannii Dendropanax.arboreus Desmopsis.panamensis
## 1                 0                 2                    0                    0
## 2                 0                 2                    3                    0
## 3                 0                 1                    6                    4
## 4                 0                 0                    0                    0
## 5                 0                 3                    5                    0
## 6                 0                 0                    2                    0
##   Diospyros.artanthifolia Dipteryx.oleifera Drypetes.standleyi Elaeis.oleifera
## 1                       1                 1                  2               0
## 2                       1                 1                  1               0
## 3                       1                 3                  2               0
## 4                       1                 0                  0               0
## 5                       0                 0                  0               0
## 6                       0                 0                  0               0
##   Enterolobium.schomburgkii Erythrina.costaricensis Erythroxylum.macrophyllum
## 1                         0                       0                         0
## 2                         0                       0                         1
## 3                         0                       0                         0
## 4                         0                       0                         0
## 5                         0                       0                         0
## 6                         0                       3                         0
##   Eugenia.florida Eugenia.galalonensis Eugenia.nesiotica Eugenia.oerstediana
## 1               0                    0                 0                   3
## 2               1                    0                 0                   2
## 3               0                    0                 1                   5
## 4               7                    0                 0                   1
## 5               2                    0                 0                   5
## 6               0                    0                 0                   2
##   Faramea.occidentalis Ficus.colubrinae Ficus.costaricana Ficus.insipida
## 1                   14                0                 0              0
## 2                   36                1                 0              0
```

```
## 3                  39               0                0             0
## 4                  39               0                0             0
## 5                  22               0                0             0
## 6                  16               0                0             0
##   Ficus.maxima Ficus.obtusifolia Ficus.popenoei Ficus.tonduzii Ficus.trigonata
## 1            1                 0              0              0               0
## 2            0                 0              0              0               0
## 3            0                 0              0              1               0
## 4            0                 0              0              2               0
## 5            0                 0              0              1               0
## 6            0                 0              0              0               0
##   Ficus.yoponensis Garcinia.intermedia Garcinia.madruno Genipa.americana
## 1                1                   0                4                0
## 2                0                   1                0                0
## 3                0                   1                0                1
## 4                0                   3                0                0
## 5                0                   2                1                0
## 6                1                   1                0                0
##   Guapira.myrtiflora Guarea.fuzzy Guarea.grandifolia Guarea.guidonia
## 1                  3            1                  0               2
## 2                  1            1                  0               6
## 3                  0            0                  0               2
## 4                  1            1                  0               5
## 5                  1            3                  0               3
## 6                  7            0                  0               4
##   Guatteria.dumetorum Guazuma.ulmifolia Guettarda.foliacea Gustavia.superba
## 1                   6                 0                  1               10
## 2                  16                 0                  5                5
## 3                   6                 0                  1                0
## 4                   3                 1                  2                1
## 5                   9                 0                  1                3
## 6                   7                 0                  0                1
##   Hampea.appendiculata Hasseltia.floribunda Heisteria.acuminata
## 1                    0                    5                   0
## 2                    0                    9                   0
## 3                    1                    4                   0
## 4                    0                   11                   0
## 5                    0                    9                   1
## 6                    0                    2                   1
##   Heisteria.concinna Hirtella.americana Hirtella.triandra Hura.crepitans
## 1                  4                  0                21              0
## 2                  5                  0                14              0
## 3                  4                  0                 5              0
## 4                  6                  0                 4              0
## 5                  4                  0                 6              0
## 6                  8                  0                 6              2
##   Hieronyma.alchorneoides Inga.acuminata Inga.cocleensis Inga.goldmanii
## 1                       0              0               2              0
## 2                       2              0               4              0
## 3                       0              0               4              1
## 4                       0              0               6              0
## 5                       0              0               0              2
## 6                       0              0               0              1
##   Inga.laurina Inga.semialata Inga.nobilis Inga.oerstediana Inga.pezizifera
```

```
## 1            0             0            0              0             0
## 2            0             0            0              0             0
## 3            0             2            1              0             0
## 4            0             4            3              0             0
## 5            1             0            1              0             0
## 6            0             0            0              0             0
##   Inga.punctata Inga.ruiziana Inga.sapindoides Inga.spectabilis
## 1             3             0                2                0
## 2             0             0                0                2
## 3             0             0                3                0
## 4             0             0                2                1
## 5             0             0                5                0
## 6             0             0                0                0
##   Inga.umbellifera Jacaranda.copaia Lacistema.aggregatum Lacmellea.panamensis
## 1                0                6                    1                    1
## 2                0               10                    0                    0
## 3                0                9                    0                    0
## 4                1                2                    1                    2
## 5                0                3                    1                    2
## 6                0                7                    2                    1
##   Laetia.procera Laetia.thamnia Lafoensia.punicifolia Licania.hypoleuca
## 1              0              0                     0                 0
## 2              1              1                     0                 0
## 3              1              1                     0                 0
## 4              0              0                     0                 0
## 5              1              0                     0                 1
## 6              0              0                     0                 0
##   Licania.platypus Lindackeria.laurina Lonchocarpus.heptaphyllus
## 1                0                   0                        7
## 2                0                   0                        7
## 3                0                   0                        3
## 4                0                   0                        9
## 5                0                   0                        2
## 6                0                   0                        1
##   Luehea.seemannii Macrocnemum.roseum Maquira.guianensis.costaricana
## 1                1                  0                             4
## 2                0                  0                             3
## 3                0                  0                             7
## 4                0                  0                             7
## 5                1                  0                            10
## 6                1                  0                             4
##   Margaritaria.nobilis Marila.laxiflora Maytenus.schippii Miconia.affinis
## 1                    0                1                 2               0
## 2                    0                0                 0               0
## 3                    0                0                 0               0
## 4                    0                0                 1               1
## 5                    1                0                 0               0
## 6                    0                0                 1               0
##   Miconia.argentea Miconia.elata Miconia.hondurensis Mosannona.garwoodii
## 1                2             0                   0                   1
## 2                0             0                   0                   0
## 3                1             0                   0                   0
## 4                0             0                   0                   0
## 5                1             0                   0                   1
```

```
## 6                0             0                   0                    1
##   Myrcia.gatunensis Myrospermum.frutescens Nectandra.cissiflora
## 1                 1                      0                    0
## 2                 0                      0                    1
## 3                 0                      0                    2
## 4                 0                      0                    2
## 5                 0                      0                    2
## 6                 0                      2                    0
##   Nectandra.lineata Nectandra.purpurea Ochroma.pyramidale Ocotea.cernua
## 1                 0                  1                  1             0
## 2                 0                  0                  0             0
## 3                 0                  0                  0             1
## 4                 0                  0                  0             1
## 5                 0                  0                  0             0
## 6                 0                  1                  0             0
##   Ocotea.oblonga Ocotea.puberula Ocotea.whitei Oenocarpus.mapora
## 1              0               0            1                22
## 2              0               0            0                21
## 3              1               0            2                14
## 4              2               2            3                23
## 5              0               0           16                17
## 6              0               1            3                19
##   Ormosia.amazonica Ormosia.coccinea Ormosia.macrocalyx Pachira.quinata
## 1                 0                0                  0              0
## 2                 0                0                  0              0
## 3                 0                0                  0              0
## 4                 0                0                  0              0
## 5                 0                0                  0              0
## 6                 0                0                  0              0
##   Pachira.sessilis Perebea.xanthochyma Cinnamomum.triplinerve
## 1                0                   0                     0
## 2                0                   1                     0
## 3                0                   0                     1
## 4                0                   0                     0
## 5                0                   1                     1
## 6                0                   0                     0
##   Picramnia.latifolia Piper.reticulatum Platymiscium.pinnatum
## 1                   0                 0                    3
## 2                   0                 0                    3
## 3                   1                 0                    5
## 4                   0                 0                    1
## 5                   0                 2                    1
## 6                   0                 0                    1
##   Platypodium.elegans Posoqueria.latifolia Poulsenia.armata Pourouma.bicolor
## 1                   2                    0              24                5
## 2                   1                    1              16                3
## 3                   3                    0              28                0
## 4                   0                    0              15                0
## 5                   0                    0              25                1
## 6                   2                    0              15                0
##   Pouteria.fossicola Pouteria.reticulata Pouteria.stipitata Prioria.copaifera
## 1                  0                   5                  0                13
## 2                  0                   7                  0                12
## 3                  0                   3                  1                12
```

```
## 4                  0                  6                  0                  5
## 5                  0                  5                  0                  3
## 6                  0                  4                  0                 26
##   Protium.costaricense Protium.panamense Protium.tenuifolium
## 1                    5                  2                  11
## 2                    4                  0                   8
## 3                    1                  2                   3
## 4                    3                  3                   9
## 5                    7                  2                   3
## 6                    1                  1                   2
##   Pseudobombax.septenatum Psidium.friedrichsthalianum Psychotria.grandis
## 1                       0                          0                   0
## 2                       0                          0                   0
## 3                       0                          0                   0
## 4                       0                          0                   0
## 5                       0                          0                   0
## 6                       0                          0                   0
##   Pterocarpus.rohrii Quararibea.asterolepis Quassia.amara Randia.armata
## 1                  1                    11             0             3
## 2                  0                    12             0             2
## 3                  0                    15             0             1
## 4                  2                    14             0             4
## 5                  1                     9             0             2
## 6                  1                     3             0             9
##   Sapium.broadleaf Sapium.glandulosum Schizolobium.parahyba Senna.dariensis
## 1                0                  0                     0              0
## 2                0                  0                     0              0
## 3                0                  1                     0              0
## 4                0                  0                     0              0
## 5                0                  2                     0              0
## 6                0                  0                     1              0
##   Simarouba.amara Siparuna.guianensis Siparuna.pauciflora Sloanea.terniflora
## 1              14                   3                    0                 1
## 2               6                   2                    0                 0
## 3              16                   1                    1                 2
## 4               8                   2                    0                 2
## 5               7                   0                    3                 3
## 6               7                   1                    0                 2
##   Socratea.exorrhiza Solanum.hayesii Sorocea.affinis Spachea.membranacea
## 1                 15               0               1                  0
## 2                 22               0               1                  0
## 3                 31               0               1                  0
## 4                  9               0               1                  0
## 5                 55               1               0                  0
## 6                 44               0               1                  0
##   Spondias.mombin Spondias.radlkoferi Sterculia.apetala
## 1               1                   2                1
## 2               1                   0                2
## 3               0                   3                0
## 4               1                   3                0
## 5               1                   5                0
## 6               0                   0                1
##   Swartzia.simplex.var.grandiflora Swartzia.simplex.continentalis
## 1                                3                              1
```

```
## 2                              3                             4
## 3                              0                             2
## 4                              1                             2
## 5                              1                             1
## 6                              9                             5
##   Symphonia.globulifera Handroanthus.guayacan Tabebuia.rosea
## 1                     0                     1              1
## 2                     1                     0              2
## 3                     1                     1              1
## 4                     1                     0              2
## 5                     2                     0              3
## 6                     0                     1              0
##   Tabernaemontana.arborea Tachigali.versicolor Talisia.nervosa Talisia.princeps
## 1                       9                    6               0                1
## 2                       5                    1               0                0
## 3                       6                    3               0                0
## 4                      10                    3               0                0
## 5                      16                    0               0                0
## 6                      11                    1               0                0
##   Terminalia.amazonia Terminalia.oblonga Tetragastris.panamensis
## 1                   1                  0                      5
## 2                   0                  0                      7
## 3                   0                  0                     10
## 4                   0                  0                     10
## 5                   1                  0                      7
## 6                   1                  0                     17
##   Tetrathylacium.johansenii Theobroma.cacao Thevetia.ahouai Tocoyena.pittieri
## 1                         0              1               0                 0
## 2                         0              1               0                 1
## 3                         0              0               0                 0
## 4                         0              0               0                 0
## 5                         0              1               0                 0
## 6                         0              0               0                 0
##   Trattinnickia.aspera Trema.micrantha Trichanthera.gigantea Trichilia.pallida
## 1                    3               0                     0                 0
## 2                    1               0                     0                 1
## 3                    1               0                     0                 0
## 4                    0               2                     0                 1
## 5                    2               1                     0                 0
## 6                    0               0                     0                 0
##   Trichilia.tuberculata Trichospermum.galeottii Triplaris.cumingiana
## 1                    18                       0                   0
## 2                    27                       0                   0
## 3                    28                       0                   0
## 4                    35                       0                   0
## 5                    15                       0                   0
## 6                    31                       0                   1
##   Trophis.caucana Trophis.racemosa Turpinia.occidentalis Unonopsis.pittieri
## 1               2                1                     0                 1
## 2               0                1                     1                 5
## 3               0                0                     1                12
## 4               0                1                     4                 3
## 5               2                0                     2                 4
## 6               0                0                     1                 3
```

```
##   Virola.multiflora Virola.sebifera Virola.surinamensis Vismia.baccifera
## 1                 0              17                   4                0
## 2                 0              12                   3                0
## 3                 0              11                   2                0
## 4                 0              16                   2                0
## 5                 0              31                   6                0
## 6                 2              19                   1                0
##   Vochysia.ferruginea Xylopia.macrantha Zanthoxylum.ekmanii
## 1                   0                 1                   3
## 2                   0                 0                   4
## 3                   0                 0                   8
## 4                   0                 0                  13
## 5                   0                 0                   3
## 6                   0                 0                   1
##   Zanthoxylum.juniperinum Zanthoxylum.panamense Zanthoxylum.setulosum
## 1                       0                     2                     0
## 2                       0                     2                     0
## 3                       1                     2                     0
## 4                       1                     5                     0
## 5                       0                     5                     0
## 6                       0                     3                     0
##   Zuelania.guidonia
## 1                 0
## 2                 0
## 3                 0
## 4                 1
## 5                 0
## 6                 2
```

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   2   3   4   5   6   7   8   9  10  11  12  13  14  15  16  17  18  19  20
##  93  84  90  94 101  85  82  88  90  94  87  84  93  98  93  93  93  89 109 100
##  21  22  23  24  25  26  27  28  29  30  31  32  33  34  35  36  37  38  39  40
##  99  91  99  95 105  91  99  85  86  97  77  88  86  92  83  92  88  82  84  80
##  41  42  43  44  45  46  47  48  49  50
## 102  87  86  81  81  86 102  91  91  93
```

```
specnumber(BCI) #they work same
```

```
##   1   2   3   4   5   6   7   8   9  10  11  12  13  14  15  16  17  18  19  20
##  93  84  90  94 101  85  82  88  90  94  87  84  93  98  93  93  93  89 109 100
##  21  22  23  24  25  26  27  28  29  30  31  32  33  34  35  36  37  38  39  40
##  99  91  99  95 105  91  99  85  86  97  77  88  86  92  83  92  88  82  84  80
##  41  42  43  44  45  46  47  48  49  50
## 102  87  86  81  81  86 102  91  91  93
```

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***: S.obs and specnumber() work same. The richness of the first four sites are 93, 84, 90, and 94 species, respectively.

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.

```
C <- function(x = ""){
  1 - (rowSums(x == 1) / rowSums(x))
}
C(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***: 0.8705882 to 0.9468504

```
C1 <- C(BCI)
range(C1)
```

```
## [1] 0.8705882 0.9468504
```

> ***Answer 2b***: Then, the coverages will be 0 and it means every species only occured once. There are so many rare species there.

Answer 2c: 0.06919643

```
C2 <- as.matrix(C1)
1 - C2[1, ]
```

```
##          1
## 0.06919643
```

Answer 2d: BCI plots' coverages range from 0.8705882 to 0.9468504, and I believe this values of coverage are pretty high.

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[1, ]
S.obs(soilbac1)
```

```
## T1_1
## 1074
```

```
C(soilbac1)
```

```
##      T1_1
## 0.6479471
```

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: 2119

Answer 3b: 1074

Answer 3c: site1's coverage is 0.9308036 which is much higher than coverage of soilbac1, 0.6479471.

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`.

```r
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
  Q1 = sum(colSums(SbyS.pa) == 1)
  Q2 = sum(colSums(SbyS.pa) == 2)
  S.chao2 = S.obs(x) + (Q1^2)/(2 * Q2)
  return(S.chao2)
}

S.ace <- function(x = "", thresh = 10){
  x <- x[x>0]
  S.abund <- length(which(x > thresh))
  S.rare <- length(which(x <= thresh))
  singlt <- length(which(x == 1))
  N.rare <- sum(x[which(x <= thresh)])
  C.ace <- 1 - (singlt/N.rare)
  i <- c(1:thresh)
  count <- function(i, y){
    length(y[y == i])
  }
  a.1 <- sapply(i, count, x)
  f.1 <- (i * (i - 1)) * a.1
  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)
}

S.chao1(BCI[1, ])
```

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

```r
S.chao2(1, BCI)
```

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

```r
S.ace(BCI[1, ])
```

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

```r
S.chao1(soilbac1)
```

```
##     T1_1
## 2628.514
```

```r
S.chao2(1, soilbac.t)
```

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

```r
S.ace(soilbac1)
```

```
## [1] 4465.983
```

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?

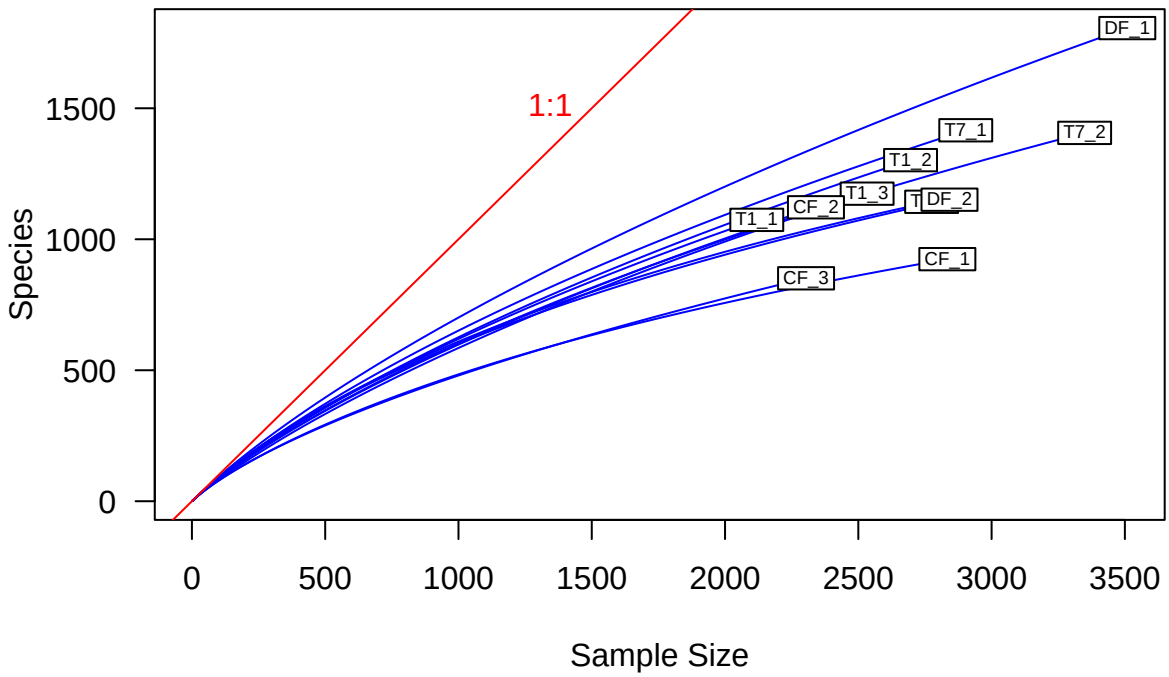
> ***Answer 4***: In contrast to Chao estimators, ACE sets the threshold of 10. Thus, we should be careful to use this estimator when we deal with samples which has many species of few individuals. I think for BCI data, these estimators gave me similar values. However, for soilbac data, the velues of chao1, 2 and ace are highly different. If I have to choose one estimators among them, I'll choose chao2 because our data include many rare species, and also have enough sites that we can obtain incidence data.

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.

```r
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)
}

site1 <- BCI[1, ]
rac <- RAC(x = site1)
print(rac)
```

```
##  [1] 25 24 22 21 18 17 15 14 14 13 13 12 12 11 11 10  9  8  7  6  6  6  6  5  5
## [26]  5  5  5  5  4  4  4  4  4  4  3  3  3  3  3  3  3  3  3  2  2  2  2  2  2
## [51]  2  2  2  2  2  2  2  2  2  2  2  2  1  1  1  1  1  1  1  1  1  1  1  1  1
## [76]  1  1  1  1  1  1  1  1  1  1  1  1  1  1  1  1  1  1
```

```
ranks <- as.vector(seq(1, length(rac)))
print(ranks)
```

```
##  [1]  1  2  3  4  5  6  7  8  9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25
## [26] 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50
## [51] 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75
## [76] 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93
```

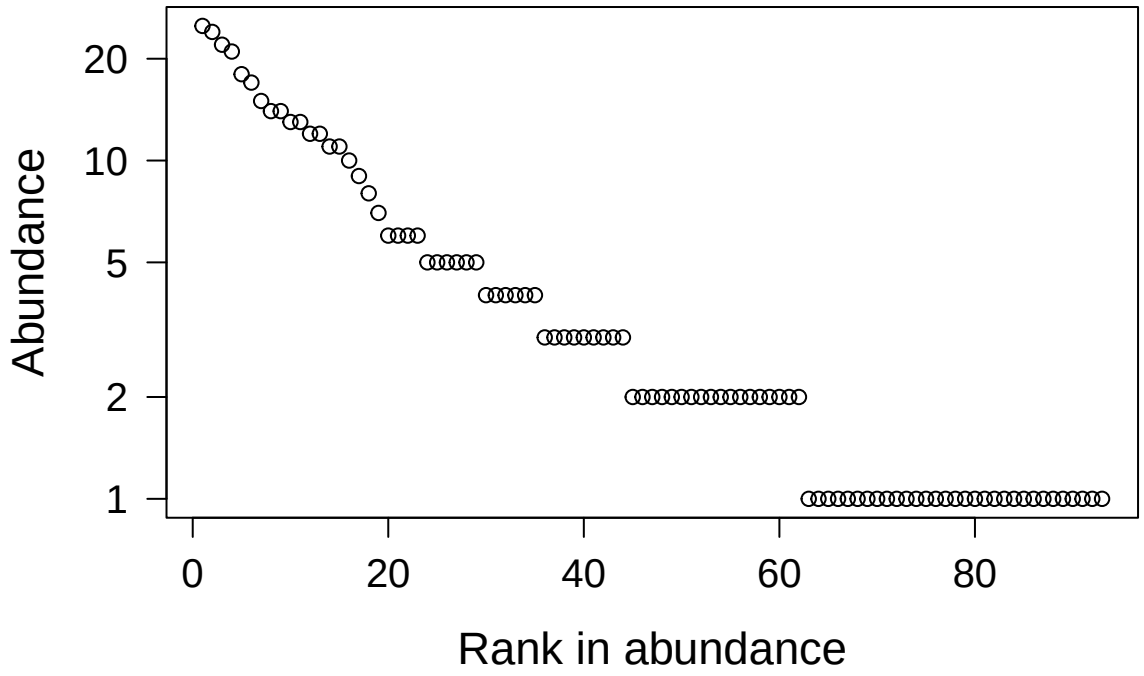
Now, let us examine the RAC for `site1` of the BCI data set.

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()
opar <- par(no.readonly = TRUE)
par(mar = c(5.1, 5.1, 4.1, 2.1))
plot(ranks, log(rac), 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)))
```



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***: I think using log-scale help us to recognize the distributino of data easily. If there is a species wich has a exceptionally high population, it makes very hard to see the distribution of other species. However, if we use log-scaled values, then we can clearly see the distribution of both dominant and rare species.

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
```

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***: Each value is 0.4238232 and 0.5067211, respectively, which I believe they are similar. For our site1 data, it doesn't have extremely dominant species, and I think that's the reason why they are not that different.

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 <- SimpD(site1)
D.inv <- 1 / D
D.sub <- 1 - D
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.

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

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

```
invD <- diversity(rac, "inv")
invD
```

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

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 alpha assumes log-series distribution which estimates diversity instead of just calculating a diversity metric, which is very useful when comparing sites with different sampling efforts.

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***:

```
D_0 = S.obs(site1)
D_1 = exp(diversity(site1, index="shannon"))
D_2 = diversity(site1, index="invsimpson")
print(paste("Hill number for q=0:", D_0))
```

```
## [1] "Hill number for q=0: 93"
```

```
print(paste("Hill number for q=1:", D_1))
```

```
## [1] "Hill number for q=1: 55.612703881371"
```

```
print(paste("Hill number for q=2:", D_2))
```

```
## [1] "Hill number for q=2: 39.4155538098979"
```

> ***Answer 8b***: I think the sharp drop in diversity from q = 0 to 2 means that there are many rare species and only a few dominant species define ecosystem function.

##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.

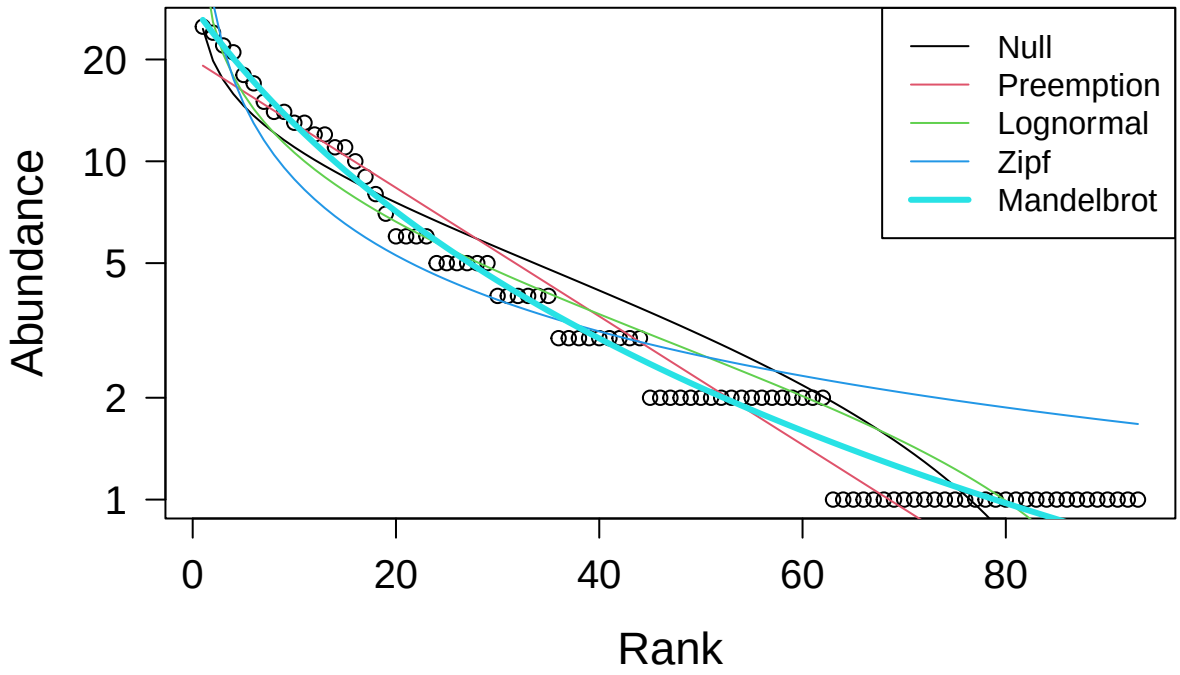
In the R code chunk below, please do the following:

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)
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***: Based on our graph, I think Mandelbrot model is the best fit for our rnaked-abundance curve. ***Answer 9b***: I think since the Manderbroy model best fit our data, it means that species abundances follow a combination of competitive dominance and stochastic processes.

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 asumes that the N is proportionla to available resource. ***Answer 10b***: RAD plot i plotted on log-cale and preemption model follows an exponential curve. Thus, it transformed into a linear matter.

Question 11: 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***: Many parameters can improve fitting data better, but we should be aware of overfitting, Thus, we can uses AIC and BIC to select the best model.

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 <- function(x = ""){
  D = 0
  N = sum(x)
  for (n_i in x){
    D = D + (n_i^2)/(N^2)
```

```
  }
  return(D)
}
D <- SimpD(site1)
D
```

```
## [1] 0.0253707
```

```
1 / D
```

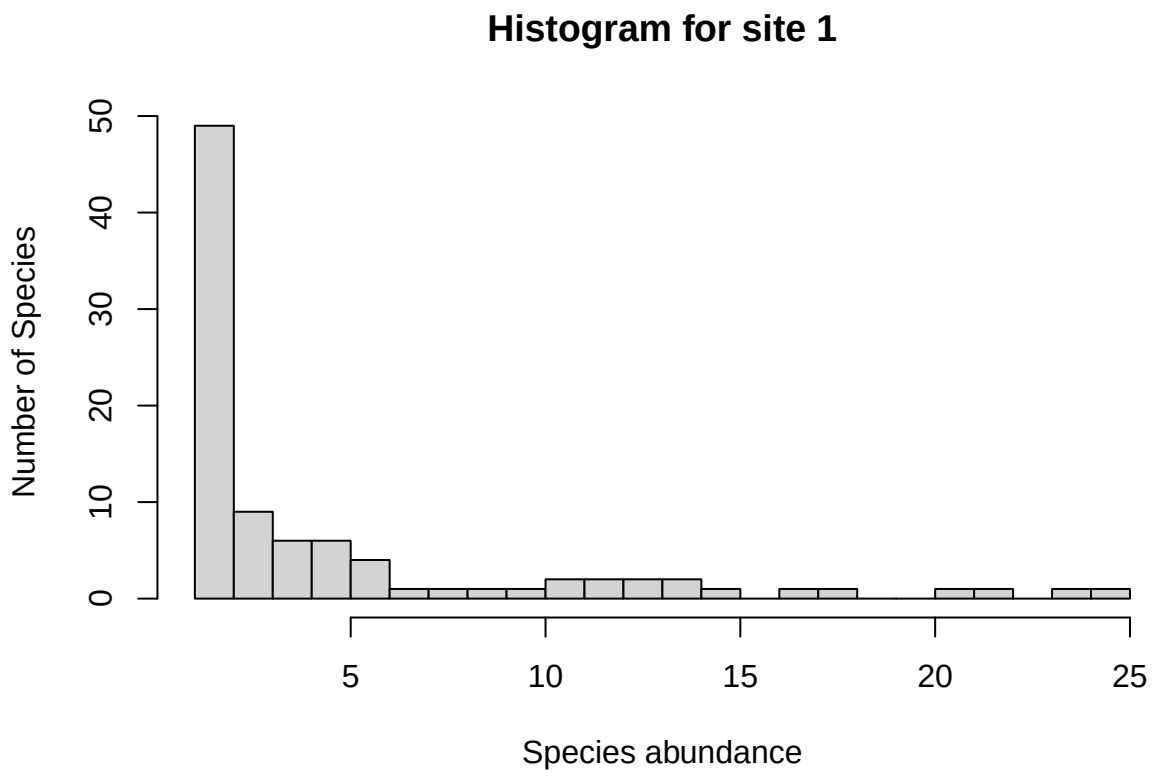
```
## [1] 39.41555
```

```
1 - D
```

```
## [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.

```
hist(rac,
     breaks = 20,
     main = "Histogram for site 1",
     xlab = "Species abundance",
     ylab = "Number of Species")
```



> Like the RAC, we can see an unequal distribution of species. About 50 species occurred only once, and the number of species drastically decreases after an abundance of 2.

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. How many sites are there? How many species are there in the entire site-by-species matrix? Any other interesting observations based on what you learned this week?

```
fish_data <- read.csv("/cloud/project/QB2025_Choi/Fish_Dataset.csv")
num_sites <- length(unique(fish_data$SiteID))
print(num_sites)
```

```
## [1] 2753
```

```
species <- fish_data[ ,23:658]
species_list <- colnames(species)
length(species_list)
```

```
## [1] 636
```

I think I haven't found some interesting observation yet. But, this data sets provide the information of divere environmental factors and species, so I believe I will find something as I make more progress.

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)**.