

# 6. Worksheet: Diversity Sampling

*Diego Rios; Z620: Quantitative Biodiversity, Indiana University*

*30 January, 2019*

## OVERVIEW

In this worksheet, you will use the jelly bean site-by-species matrix generated from **6. Diversity Sampling**. Along with tools outlined in the **5. Local (alpha) Diversity** and **7. Control Structures** handouts, you will develop questions, visualize data, and test hypotheses related to sampling effects and its effect on estimates of within-sample biodiversity.

## 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. Refer to previous handout to help with developing of questions and writing of code.
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 `6.DiversitySampling_Worskheet.Rmd` and the PDF output of Knitr (`DiversitySampling_Worskheet.pdf`).

## 1) Group brainstorming

With your team partner and perhaps other students in the class, spend 15 minutes or so brainstorming questions, code, “fantasy figures”, and statistical tests that could be used to test questions with the class’s data represented in the site-by-species matrix that you have generated.

## 2) Code

Use the space below for code that is being used to analyze your data and test your hypotheses. Create one (and only one, although it can have multiple panels) *publication quality* figure. Make sure to annotate your code using # symbols so others (including instructors) understand what you have done and why you have done it.

Load the data

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

```
## function ()
## .Internal(getwd())
## <bytecode: 0x7fd31d40f578>
```

```
## <environment: namespace:base>
```

```
setwd("~/GitHub/QB2019_Rios/2.Worksheets/6.DiversitySampling/")
JB <- read.delim("JellyBeans.txt", header = T)
```

Perform statistics

```
#Coverage
C <- function(x = ""){
  1 - (sum(x == 1) / sum(x))
}

#Species richness
S.obs <- function(x = ""){
  rowSums(x > 0) * 1
}

#Chao 1 index
S.chao1 <- function(x = ""){
  S.obs(x) + (sum(x == 1)^2) / (2 * sum(x == 2))
}

#ACE index
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) # counter to go through i range
  count <- function(i, y) {
    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)
}

Coverage <- apply(JB[, -c(1,2)], 1, C)
S <- S.obs(JB[, -c(1,2)])
Chao1 <- S.chao1(JB[, -c(1,2)])
ACE <- apply(JB[, -c(1,2)], 1, S.ace)

Group <- c("A", "A", "A", "B", "B", "A", "B", "A", "B")
JB1 <- as.data.frame(cbind(Group, Coverage, S, Chao1, ACE))
JB1$Coverage <- as.numeric(as.character(JB1$Coverage))
JB1$S <- as.numeric(as.character(JB1$S))
JB1$Chao1 <- as.numeric(as.character(JB1$Chao1))
JB1$ACE <- as.numeric(as.character(JB1$ACE))

library(ggplot2)
fig1a <- ggplot(JB1, aes(Group, Coverage)) + geom_boxplot(size=0.5) +
```

```

xlab("Group")+ylab("Coverage")+
  theme(axis.title.x=element_text(size=10),
        axis.text.x=element_text(size=9,color="black"),
        axis.title.y=element_text(size=10),
        axis.text.y=element_text(size=9,color="black"),
        axis.line=element_line(size=.4),
        panel.background=element_blank())

fig1b <-ggplot(JB1, aes(x=Group,y=S))+geom_boxplot(size=0.5)+
  xlab("Group")+ylab("Jelly Bean Richness")+
  theme(axis.title.x=element_text(size=10),
        axis.text.x=element_text(size=9,color="black"),
        axis.title.y=element_text(size=10),
        axis.text.y=element_text(size=9,color="black"),
        axis.line=element_line(size=.4),
        panel.background=element_blank())

fig1c <-ggplot(JB1, aes(Group,Chao1))+geom_boxplot(size=0.5)+
  xlab("Group")+ylab("Chao 1")+
  theme(axis.title.x=element_text(size=10),
        axis.text.x=element_text(size=9,color="black"),
        axis.title.y=element_text(size=10),
        axis.text.y=element_text(size=9,color="black"),
        axis.line=element_line(size=.4),
        panel.background=element_blank())

fig1d <-ggplot(JB1, aes(Group,ACE))+geom_boxplot(size=0.5)+
  xlab("Group")+ylab("ACE index")+
  theme(axis.title.x=element_text(size=10),
        axis.text.x=element_text(size=9,color="black"),
        axis.title.y=element_text(size=10),
        axis.text.y=element_text(size=9,color="black"),
        axis.line=element_line(size=.4),
        panel.background=element_blank())

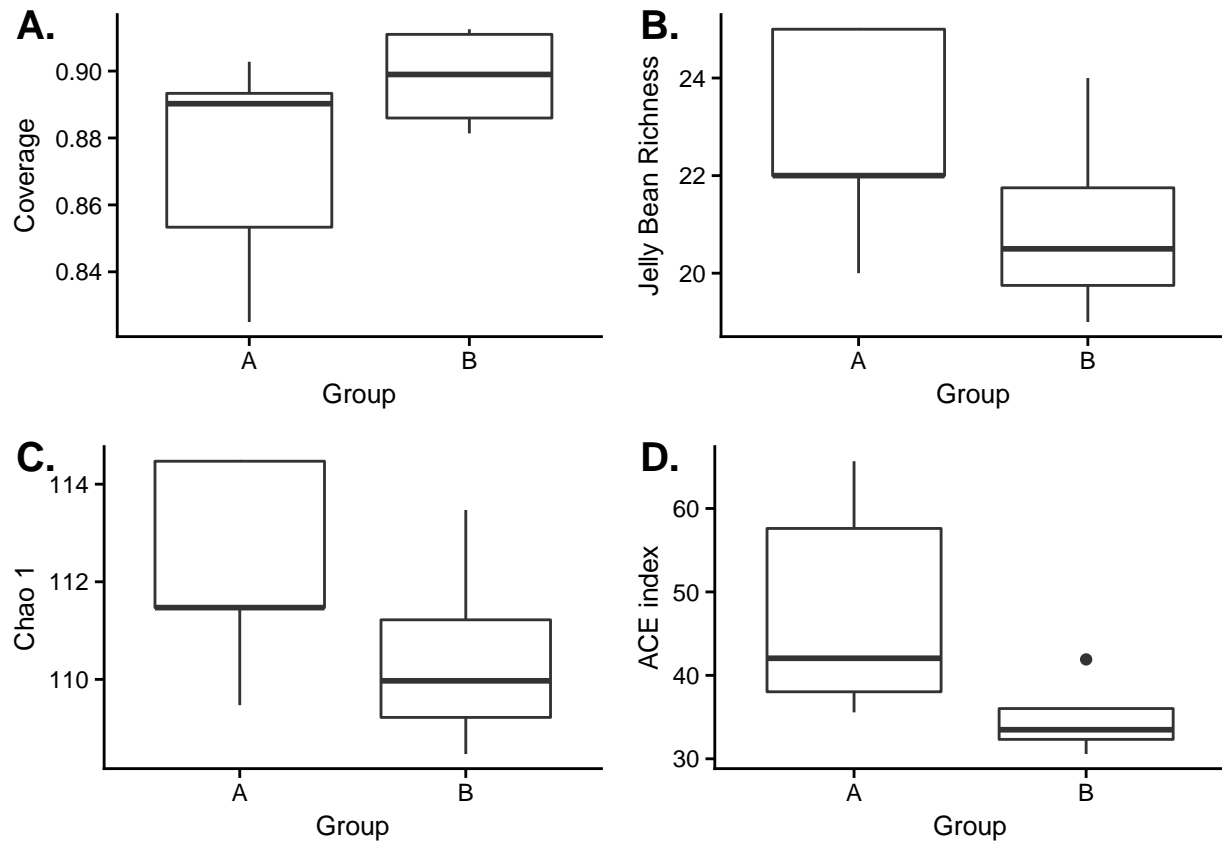
library(cowplot)

##
## Attaching package: 'cowplot'

## The following object is masked from 'package:ggplot2':
##
##      ggsave

fig1 <- plot_grid(fig1a, fig1b, fig1c, fig1d, nrow = 2,ncol=2)+
  draw_plot_label(c("A.", "B.", "C.", "D."), c(0, 0.5, 0, 0.5), c(1, 1, 0.5, 0.5), size = 15)
fig1

```



### 3) Figure caption

Write an informative yet succinct (~5 sentences) caption that creates a “stand-alone” figure. Take a peek at figures and figure captions in a paper published in your favorite journal for inspiration.

Fig. 1. Local Jelly Bean richness boxplots for two groups of emerging scientists: Group A ( $n = 5$ ) and Group B ( $n = 4$ ). (A) Good’s Coverage index of sampling effort, (B) Observed Jelly Bean Richness, (C) Chao 1 Jelly Bean Richness Estimate, and (D) ACE Jelly Bean Richness Estimate. Coverage was higher and varied less for group B, than for group A. Furthermore, the three estimates of Jelly Bean richness were consistently lower for Group B than for Group A.

## SUBMITTING YOUR ASSIGNMENT

Use Knitr to create a PDF of your completed 6.DiversitySampling\_Worksheet.Rmd document, push it to GitHub, and create a pull request. Please make sure your updated repo includes both the pdf and RMarkdown files.

Unless otherwise noted, this assignment is due on **Wednesday, January 30<sup>th</sup>, 2017 at 12:00 PM (noon)**.